Federal Agency	Department of Energy, Office of Energy Efficiency and Renewable Energy
Nature of Report	Final Technical Report
Award Number	DE-EE0007883
Award Type	Cooperative Agreement
Recipient Name	Renewable Energy Wildlife Institute (formerly American Wind Wildlife Institute): Lauren Flinn
Recipient Title	Director of Operations
Recipient Email	lflinn@rewi.org
Recipient Phone Number	(202) 448-8780
Prime Recipient Type	Non-Profit
Project Title	Evaluating the Effectiveness of a Detection and Deterrent System in Reducing Golden Eagle Fatalities at Operational Wind Facilities
Principle Investigator	Shilo Felton, sfelton@rewi.org
Recipient DUNS Number	015077939
Date of Report	May 31 2024
Period Covered by Report	June 1, 2017 – May 31, 2024
Attachments	1. Statement of Project Objectives (MOD 10) 2. Budget Information (MOD 12) 3. Milestone 8.2 Report: Goodnoe Hills Two-year Experiment 4. Milestone 9.1 Report: Goodnoe Hills Eagle Behavioral Responses to Deterrents 5. Milestone 10.1 Report: Multi-site Analysis of DTBird Detection and Deterrence Trigger Distances 6. Milestone 10.2 Report: Multi-site Analysis of False Positive and False Negative DTBird Detection Rates 7. Milestone 10.3 Report: Multi-site Analysis of Avian Behavioral Response to Deterrents 8. Milestone 10.4 Report: Multi-site Estimate of Eagle Collision Risk Reduction 9. Milestone 11.1 Report: DTBird Cost Analysis

Evaluating the Effectiveness of a Detection and Deterrent System in Reducing Golden Eagle Fatalities at Operational Wind Facilities

Award Number: DE-EE0007883

May 31, 2024

Renewable Energy Wildlife Institute (REWI)

700 12th Street, NW, Suite 700

Washington, DC 20005

www.rewi.org

Prepared By

Shilo K. Felton, Ph.D., Senior Scientist, REWI

La' Portia J. Perkins, M.Sc., Project Manager, REWI

Jeff P. Smith, Ph.D., Associate Wildlife Ecologist, H. T. Harvey & Associates

Scott B. Terrell, Ph.D., Principal Wildlife Ecologist, H. T. Harvey & Associates

Contents

Section 1. Executive Summary

The Renewable Energy Wildlife Institute (REWI) was appointed as the prime awardee of DOE award number DE-EE0007883 to lead a team of scientists, wind developers, and technology manufacturers toward the overarching goal of evaluating the effectiveness of the current DTBird system in minimizing the risk of golden eagles (*Aquila chrysaetos*) and other large soaring raptors from approaching the rotor-swept zone (RSZ) of operating wind turbines. As part of this goal, the team set out to 1) quantify the expected reduction in collision risk for golden eagles from operation of the detection and deterrence modules in a manner that supports the approach used by the U.S. Fish and Wildlife Service (USFWS) to assess and credit facility operators for their efforts to minimize predicted collision fatalities and 2) provide information to help improve the technology to maximize its effectiveness.

DTBird is an automated detection and audio deterrent system created by the Spanish company Liquen, designed to discourage birds from entering the RSZ of spinning wind turbines. The system uses cameras to automatically detect airborne targets of interest, records each such event in an online database, and triggers a warning signal (loud sound) if the tracked object has moved close to the turbine. If the object moves even closer to the RSZ, a more aggressive dissuasion signal is broadcast.

To meet our objectives, the team conducted a two-year experiment at the Goodnoe Hills wind facility in Washington state, in which 14 turbines were outfitted with DTBird units. Daily, each DTBird-equipped turbine was randomly assigned to a control or treatment group. Treatment turbines operated with DTBird running as intended—broadcasting warning or deterrent signals when DTBird detected a target within range. On control turbines, no sound signals were broadcast if a moving target triggered the DTBird system. The team also flew unmanned aerial vehicles (UAVs) designed to coarsely mimic the general size, weight, and coloration of golden eagles in programmed flight transects across DTBird detection ranges to quantify DTBird's ability to detect intended targets and to evaluate factors that influence the probability of detection and DTBird's response distances. Additionally, the team evaluated the behavioral responses of *in situ* eagles exposed to spinning turbines alone (visual and sound influences) versus spinning turbines plus broadcasted DTBird audio deterrents, to estimate the effectiveness of deterrence by the DTBird system. The data and results from these investigations were combined with those from a pilot study conducted at the Manzana Wind Power Project in California to better evaluate DTBird's effectiveness across different landscapes.

Results of the controlled two-year experiment in Washington indicated that broadcasted deterrence signals significantly reduced the time eagles spent near DTBird-equipped turbines (aka dwell time). However, the initial warning signals did not significantly influence the rate at which eagles triggered more intense dissuasion signals, likely because eagles often entered the dissuasion signal zone without first being detected by DTBird within the warning signal zone. There was also a strong interactive effect of deterrence-signal and false-positive rates, meaning that if warning/dissuasion signals were triggered and broadcast more frequently at experimental turbines due to false positives (e.g., detection events triggered by birds other than

eagles, or by non-bird objects like rotor blades, airplanes, or clouds) eagles generally spent less time around those turbines. These results suggest that, despite our concerns that high false-positive rates could cause eagles to become less responsive to deterrent signals, negative habituation did not occur over this experiment. Additionally, although the video quality made it difficult to confidently classify events as successful, 53–100% of all probable golden and bald eagles exhibited a successful (or potentially successful) response to deterrent and warning signals, rates that matched or exceeded the established performance metric of ≥50% successful deterrence for eagles.

Overall, the behavioral results from both the California and Washington wind facilities indicated that operation of DTBird reduced the overall likelihood that an eagle passing through the expected detection range would approach the RSZ by 20–30%, and that value increased by at least 5–10% for birds classified as at moderate to high risk of approaching the RSZ prior to deterrent signaling (moderate risk: bird on a course taking it near but not directly toward the RSZ; high risk: bird on course to intersect with RSZ). Both multi-species and golden eagle analyses confirmed response differences at the two facilities and in relation to preexposure risk levels, and the multi-species model also emphasized that species responded differently to wind speed. The probability of effective deterrence was generally highest for birds classified as at moderate preexposure risk of approaching the RSZ, potentially because those birds had more time and space to effectively respond to the deterrents than birds making high-risk movements toward the RSZ.

Trials using eagle-like UAVs to evaluate DTBird's ability to detect golden eagles and other large raptors revealed an overall 65% probability of detection within 240 meters of the cameras, with the highest chance of detection when the target flew within 80–160 meters of the turbine versus closer or farther away. Cloudy skies, wind speed, different UAV models (potentially reflecting differences between eagle sexes and age classes), UAV speed, and pitch and roll angles all influenced the distance at which DTBird detected the UAVs.

Initially, DTBird registered non-bird objects (e.g., turbine blades, planes, shadows) relatively frequently, so Liquen adjusted the algorithms in January 2023 (5 months into Year 2 of the experiment) to lower the rate of these false positive events. Doing so lowered the false positive rate from 3.9 to 0.8 false-positive deterrence triggers/turbine/day to meet or fall under the established performance metric (1.6–2.8 triggers/turbine/day). Overall, results emphasized the value of running the detection algorithm for an additional three months prior to considering DTBird fully commissioned in the field and suited to operation with deterrents broadcasting.

The standard DTBird V4D8 model sale cost (with Falco cameras and Larus software) is around $18–$22K, and the yearly service sale cost is around $2–3K. Additionally, installation costs $4–6K/unit and maintenance runs $0.6–2K/unit/year. This brings the total investment to purchase and operate a single DTBird unit for the first year to a minimum of $24,600, based on the cost of installing, operating, and maintaining the 14 DTBird units at Goodnoe Hills.

While the cost per unit may be less than other commercially available risk-reduction systems, our study results revealed some areas in which the technology could be improved. We did not find that eagles showed negative habituation to the overactive triggers, however, we would need

further study to confirm this, and there may be other negative consequences of excessive deterrence signals. Given this better camera resolution and refined AI algorithms could greatly improve the functionality of the system and better enable target detections against various backdrops. Furthermore, we recommend users ensure regular replacement of camera lenses to avoid solar degradation which further affects target detection.

.

This report provides details on the study design and implementation as well as the tasks, milestones, costs, and challenges related to the evaluation of the DTBird system.

Section 2. Acknowledgments

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Wind Energy Technologies Office Award Number DE-EE0007883. Avangrid Renewables, Puget Sound Energy (PSE), Pacific Wind Lessee, Portland General Electric (PGE), PacifiCorp, Liquen, and H. T. Harvey & Associates provided additional support for this project. We also thank reviewers C. Hein, M. Huso, J. Garvin, and N. Lewandowski for their valuable feedback which greatly improved the study design, analyses, and this report as a whole.

Section 3. Disclaimer

Section 4. Study Narrative

4.1 Introduction

The Renewable Energy Wildlife Institute (REWI) was appointed as the prime awardee of DOE award number DE-EE0007883 to lead a team of scientists, wind developers, and technology manufacturers toward the overarching goal of evaluating the effectiveness of the current DTBird system in minimizing the risk of golden eagles (*Aquila chrysaetos*) and other large soaring raptors from approaching the rotor-swept zone (RSZ) of operating wind turbines.

Eagle collisions with wind turbines (Hunt 2002, Erickson et al. 2005, de Lucas et al. 2008, Smallwood 2013) are well-documented. The Migratory Bird Treaty Act (16 U.S.C. §703), the California Department of Fish and Game Code (§3503 and §3511), and the Bald and Golden Eagle Protection Act (Eagle Act) (16 U.S.C. §668–668c) protect eagles from human-related mortality and disturbance sufficient to cause a decline in eagle survival or productivity. We propose to evaluate the effectiveness of DTBird detection and deterrence system in reducing golden eagle collisions at wind turbines.

The DTBird system to be evaluated includes a video-surveillance detection module and a collision-avoidance or deterrence module. Technical specifications for this project are based on the model DTBirdV4D8, with four HD wide-angle cameras located every 90° on the mast of the wind turbine below the blades, and eight speakers, four located by the cameras and four at 90º on the mast of the wind turbine below the nacelle. Overlapping detection areas and improvement in sound distribution will provide good performance in detection and collision avoidance.

Previous European evaluations of DTBird provided preliminary insight into its ability to detect and deter raptors and other birds from approaching turbines (May et al. 2012, Hanagasioglu et al. 2015). Those researchers accomplished this objective primarily by comparing the frequency and turbine-approach distances of *in situ* raptors that they visually observed flying near turbines with and without the DTBird warning and dissuasion signals muted.

May et al. (2012) evaluated the ability of DTBird to detect and deter raptors flying near and in the risk zone of wind turbines in Norway, with the DTBird system calibrated to detect and deter large raptors such as white-tailed sea eagles (*Haliaeetus albicilla*) and golden eagles. The authors of this study compared the detection rates of the DTBird camera and video surveillance system against detections documented by a radar system. Using this approach, they were able to quantify false positives (i.e., video recording stimulated by activity other than target birds) and false negatives (i.e., the detection system failed to trigger video surveillance when radar indicated a target passed by in detectable range) detection rates. This study, as well as Hanagasioglu et al. 2015, did not explicitly address the potential existence and importance of "blind spots" in the DTBird detection system, nor did it evaluate detectability as a function of covariates that can only be addressed by controlled experiments using flying objects manipulated to fly under specified conditions and in predefined patterns.

In alignment with the recommendations of May et al. (2012), we used unmanned aerial vehicles (UAVs) in experimental trials to further evaluate the performance of the DTBird detection and deterrence system at two operational wind-energy facilities in distinctly different landscape settings where golden eagles occur. We will combine insight gained from these trials with data on the detection and deterrence responses of *in situ* golden eagles and large buteos, such as red-tailed hawks, recorded by the DTBird system, as well as with insight derived from a controlled field experiment designed to quantify the degree to which operation of the DTBird detection and deterrence system reduces the probability of eagles and surrogate raptors entering the collision risk zone of equipped turbines.

The primary outcomes of the study will be:

1. Statistically robust understanding of the current DTBird detection and deterrence system's ability to successfully deter golden eagles and suitable surrogate raptors from entering the collision risk zone of turbines.
2. Statistically robust understanding of the limitations of the system and how various environmental and behavioral variables influence the system's effectiveness.
3. Assistance to DTBird engineers with refining the system to maximize its performance and effectiveness in reducing conflict between eagles/raptors and operating wind turbines.
4. Effective, statistically robust projections concerning the ability of DTBird to reduce fatality rates for golden eagles and other raptors at wind facilities similar to those involved in the study, thereby assisting the USFWS, state regulatory agencies, and facility operators with projecting the anticipated risk-reduction benefits of deploying the DTBird system.

4.2 Study Objectives

To achieve the primary outcomes above and the overall goal of providing a rigorous multi-site evaluation of DTBird's ability to reduce the risk of eagles and other medium/large raptors entering the collision risk zone of operational wind turbines, we pursued the following primary objectives:

Objective 1: Quantify the probability of detection and evaluate the accuracy, precision, and limitations of the DTBird detection and deterrent-triggering functions using UAVs designed to resemble golden eagles at the California (pilot study) and Washington study sites.

Objective 2: Quantify the probability of deterrence and evaluate the effectiveness of the DTBird deterrent signals in reducing raptor activity within the RSZ of turbines by evaluating the behavioral responses of *in situ* eagles and other raptors exposed to the deterrent signals as revealed in DTBird video records from both study sites.

Objective 3: Quantify the prevalence and describe the nature of false positive detections (i.e., DTBird detections of birds and other objects that are not target raptors) at the two study sites.

Objective 4: Quantify DTBird's proximate and longer-term effectiveness in reducing eagle and surrogate raptor activity around equipped turbines by conducting a two-phase 2-year controlled experiment at the Washington study site to evaluate the comparative DTBird event triggering rates at turbines with and without muted deterrent signals.

Objective 5: Produce a multi-site estimate of the potential for DTBird to reduce the risk of eagles entering the collision risk zone of operational turbines at facilities similar to those involved in the study.

Objective 6: Evaluate the performance reliability and maintenance requirements of the DTBird systems installed at the study sites and prepare a detailed systems cost analysis.

4.3 Field Methods and Data Processing

To meet our objectives, the team conducted a two-year experiment at the Goodnoe Hills wind facility in Washington state, in which 14 turbines were outfitted with DTBird units. Daily, each DTBird-equipped turbine was randomly assigned to a control or treatment group. Treatment turbines operated with DTBird running as intended—broadcasting warning or deterrent signals when DTBird detected a target within range. On control turbines, no sound signals were broadcast if a moving target triggered the DTBird system. The team also flew UAVs designed to coarsely mimic the general size, weight, and coloration of golden eagles in programmed flight transects across DTBird detection ranges to quantify DTBird's ability to detect intended targets and to evaluate factors that influence the probability of detection and DTBird's response distances. Additionally, the team evaluated the behavioral responses of *in situ* eagles exposed to spinning turbines alone (visual and sound influences) versus spinning turbines plus broadcasted DTBird audio deterrents, to estimate the effectiveness of deterrence by the DTBird system. The data and results from these investigations were combined with those from a pilot study conducted at the Manzana Wind Power Project in California to better evaluate DTBird's effectiveness across different landscapes.

4.3.1 Study Sites

4.3.1.1 Manzana Wind Project, California

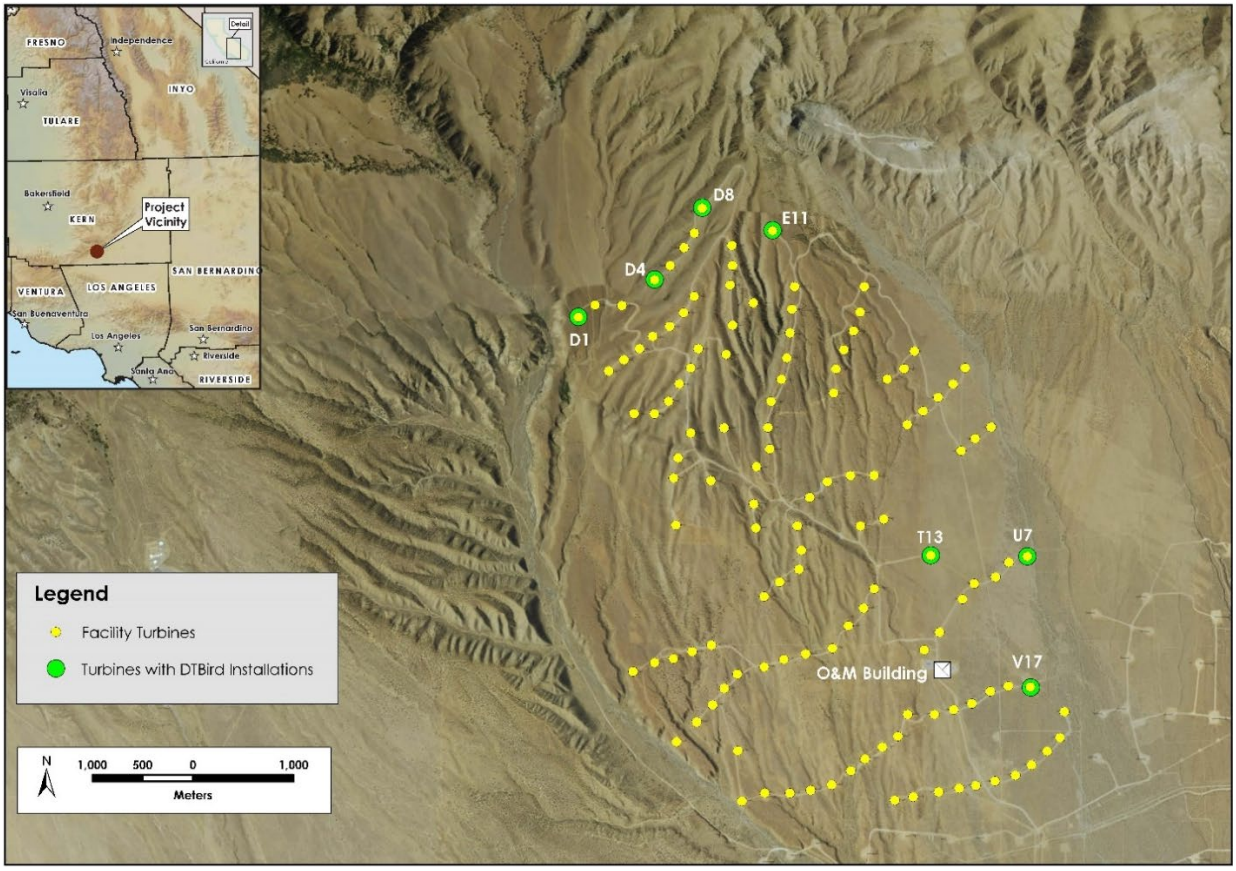

Figure 1. Layout of the Manzana Wind Power Project in southern California showing locations of installed DTBird systems.

The Manzana Wind Project has been in operation since 2012 and comprises 126 1.5 MW GE 1.5-77 wind turbines, with a hub height of 65 meters and a rotor-swept diameter of 82.5 meters, located in the southwestern foothills of the Tehachapi Mountains of southern California in northwestern Antelope Valley, which constitutes the westernmost extension of the Mojave Desert (Figure 1). The landscape is a gradually sloping alluvial fan incised by dry desert washes. The northwestern sector of the facility features more complex foothill topography adjacent to a primary riparian drainage, and the topography grades downslope to the southeast into a more-uniform plain. The desert scrub and woodland vegetation is typical of the upper Mojave Desert region. Seven DTBird systems were strategically installed here to support this research (Figure 1; H. T. Harvey & Associates 2018).

4.3.1.2 Goodnoe Hills Wind Farm, Washington

The Goodnoe Hills Wind Farm has been in operation since 2008 and currently comprises 47 2.2 MW Vestas V110 Mark C and B wind turbines, with a hub height of 87 meters and a rotor-swept diameter of 110 meters located in south-central Washington atop an east-west ridgeline

flanking the Columbia River approximately 3–6 km away (Figure 2). The topography descends steeply south of the ridgeline approximately 610 meters to the Columbia River and more

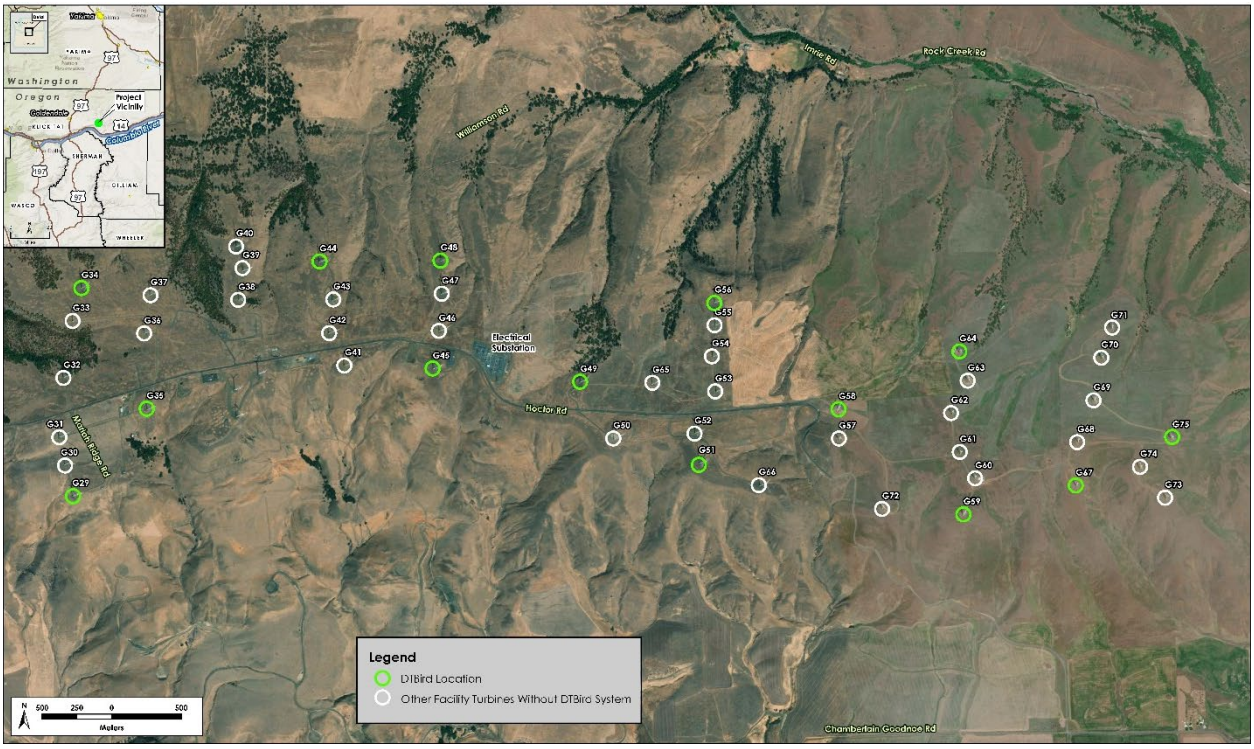

Figure 2. Layout of the Goodnoe Hills Wind Farm in south-central Washington showing locations of installed DTBird systems.

gradually to the north approximately 500 meters down into Rock Creek Canyon and associated riparian corridors. The project area is dominated by a mosaic of grazed grassland and shrubsteppe, with inclusions of ponderosa pine (Pinus ponderosa) and Oregon white oak (*Quercus garryana*) woodlands on the ridge's north-facing slopes. Fourteen DTBird systems were installed around the perimeter of this facility to support this research; however, the extent of effective operation varied among the installed systems during the 2-year study at this site (Attachment 4).

4.3.2 DTBird System Operation

DTBird is an automated detection and audio deterrent system created by the Spanish company Liquen, designed to discourage birds from entering the RSZ of spinning wind turbines. The system uses cameras to automatically detect airborne targets of interest, records each such event in an online database, and triggers a warning signal (loud sound) if the tracked object has moved close to the turbine. If the object moves even closer to the RSZ, a more aggressive dissuasion signal is broadcast.

The DTBird systems were set up with four 6-megapixel HD cameras arrayed in approximate cardinal directions on the turbine towers at a height of 4 m agl, and four speakers arrayed in similar fashion around the tower at a height close to the lower RSZ. The Goodnoe Hills

installations included a second ring of four broadcast speakers installed on the turbine towers just below hub height (Figure 3). This modification was implemented to account for taller turbines at the Goodnoe Hills and thereby help to ensure effective deterrent broadcasting throughout a larger overall detection envelope and collision risk zone. Field measurements correlated with known assigned camera numbers confirmed that the orientation of cameras of a given number was variable but nonetheless coarsely consistent across the seven installations. Camera 1 always faced to the west, Camera 2 to the south, Camera 3 to the east, and Camera 4 to the north. The systems included a light monitor that restricted their operation to periods when the lighting exceeded 50 lux, which translates to operation from civil dawn to civil twilight. In addition, during normal operations, the collision-avoidance module (deterrent signals) operated only when the turbine blades were spinning at a rate of ≥3 rpm. At the minimum cut-in wind speed for turbines at the study site (3.5 m/second [sec]), the blade rotors spun at a rate of approximately 12–14 rpm.

The broadcast volume of the deterrent signals can be adjusted depending on site-specific needs pertaining to the targeted bird species, local noise-management ordinances, and the specific facility layout. The factory setting broadcasts sounds at approximately 121 decibels (dB) at 1 m from the turbine. Sound-attenuation models and testing by Liquen during installation of the systems confirmed that broadcasting at the factory setting would not exceed the Kern County noise-ordinance restriction of ≤65 dB at the exterior of the residence closest to a DTBird installation (approximately 0.5 km). On days when UAV flight trials occurred, deterrent signals were muted at the focal turbine during all daylight hours. This arrangement was necessary to allow the operations team to maintain clear verbal communication at all times, and because the

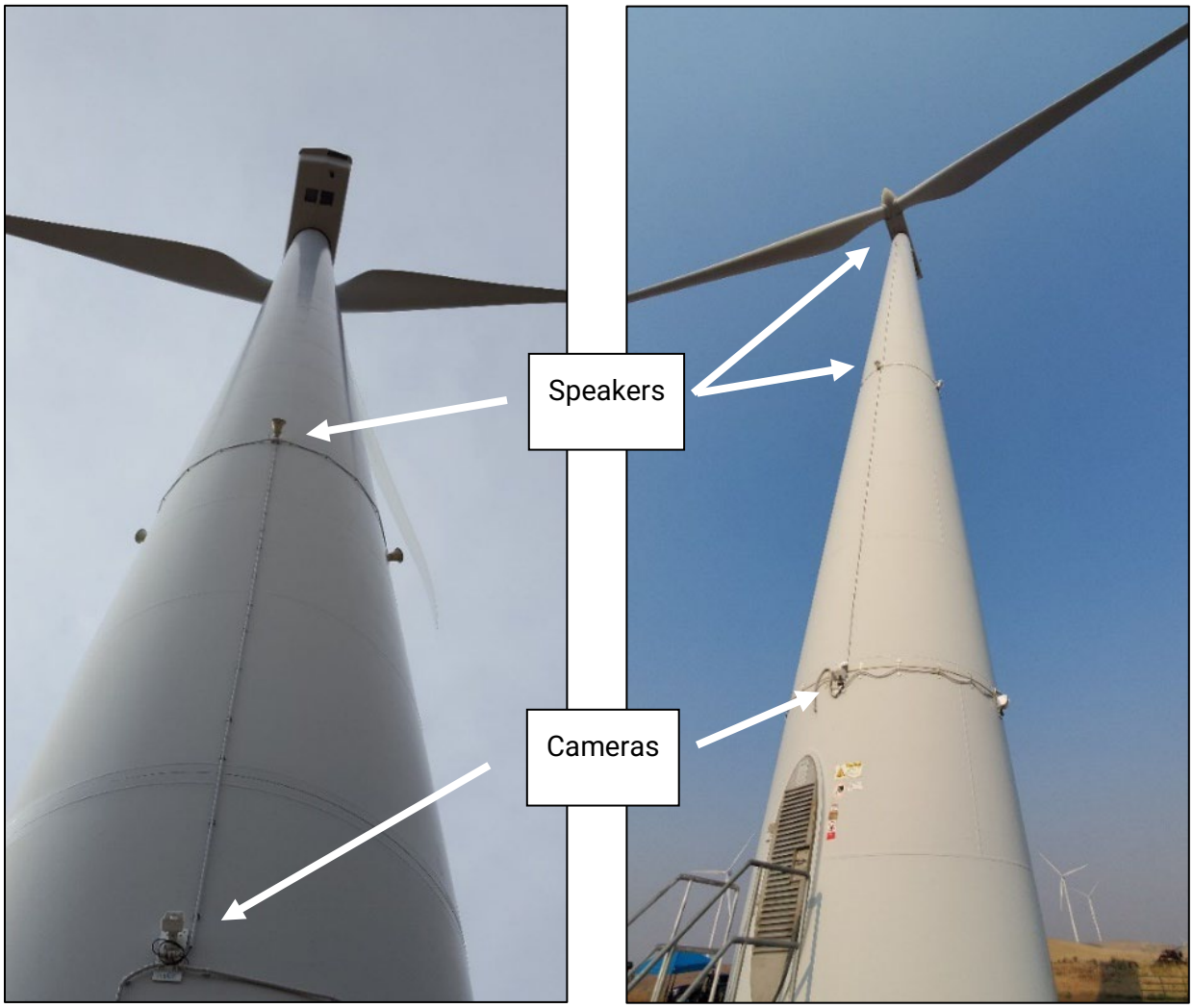

Figure 3. Depiction of DTBird video camera and broadcast speaker locations on turbines at the Manzana Wind Power Project (left panel, single ring of speakers) and Goodnoe Hills Wind Farm (right panel, two rings of speakers).

local time difference between the study site in the United States and the DTBird control operation in Spain precluded timelier coordination during the actual trials.

Each individual DTBird automated video system surveilled the sky around an individual turbine for moving objects that filled enough image pixels to qualify as a target of interest based on calibrations for the focal species of interest, in this case golden eagle. DTBird does not classify or enumerate targets, may target multiple objects simultaneously, and does not actually track individual objects—it simply repeatedly registers individual objects as targeted as long as they meet the calibrated targeting criteria. Analysts must subsequently review event records and video clips stored in the DAP to classify and enumerate the detected targets, which may be birds or false positive detections caused by airplanes, insects, debris, raindrops, snowflakes, or other inanimate objects moving through the detection envelope, as well as from by sky artifacts

(e.g., high-contrast, shifting elements caused by clouds and bright skies that are mistaken for flying objects).

DTBird systems are calibrated to target objects of a specified size range and, if a system registers that the turbine rotor is actively spinning at ≥2 rotations per minute (rpm) to trigger subsequent deterrent signals when the system estimates that a targeted object as within a specified distance from the turbine. Detection and trigger distances are determined based on pre-programmed criteria projecting how many image pixels a bird of the specified size is expected to fill at specified distances. The Manzana and Goodnoe Hills systems were calibrated to target golden eagles (wing span of 2.1–2.3 m), which translated to targeting objects that met specified criteria at an expected maximum line-of-sight distance from the turbine of approximately 240 m. Once an object is targeted and a new detection record initiated at a spinning turbine, the system triggers an initial audible warning signal if it perceives that a targeted object moves within 170–240 m of the turbine, and triggers a more aggressive dissuasion signal at distances of 100–170 m, depending on the flight altitude (Figure 4; and see H. T. Harvey & Associates 2018 for additional graphical illustrations and detailed information about the expected deterrent-triggering zones within the projected overall detection envelope).

When a DTBird system first detects a targeted object, it creates a new event record in the online digital analysis platform (DAP) database Liquen maintains to store detection records and extracted video clips for all DTBird installations. The DAP records a timestamp for each initial detection event along with other limited data. Other data automatically recorded in the DAP for each detection event include: (a) the average wind speed, rotor azimuth, and rotor rpm during the event record derived from the turbine SCADA system; (b) a binary indicator of whether or not the focal rotor was spinning sufficiently for DTBird deterrence module to be operating; (c) an estimate of the current amount of ambient illumination; and (d) length of the video tracking record. If a targeted object subsequently or simultaneously triggers one or both of two deterrent signals (early warning or a more raucous dissuasion signal if a target approaches closer to the turbine) information is added to the same DAP event record to document the unique timestamps and signal durations for each deterrent-triggering event. Each event record has video clips attached to it representing the four cameras, which the system extracts to begin 10 seconds before targeting began and continue for 30 seconds after the last targeted object exits the detection envelope. There must be no objects targeted for at least 26 seconds before a given DTBird system can initiate a new detection event record. If a system targets multiple objects concurrently during the same event period, timestamps are recorded only for the first detection, warning-trigger, and dissuasion-trigger events, and those respective events may not be triggered by the same object. In these cases, sometimes it can be difficult to determine exactly which bird or object was responsible for the timestamped events. Technicians must screen all relevant DAP records and videos to classify and enumerate the detected objects, which can include birds of all types and sizes as well as myriad other animate and inanimate flying objects, and to identify other sources of false positive detections caused by the detection system perceiving dynamic, high-contrast elements in the viewshed associated with moving turbine blades, clouds, and other turbine equipment as moving objects of interest.

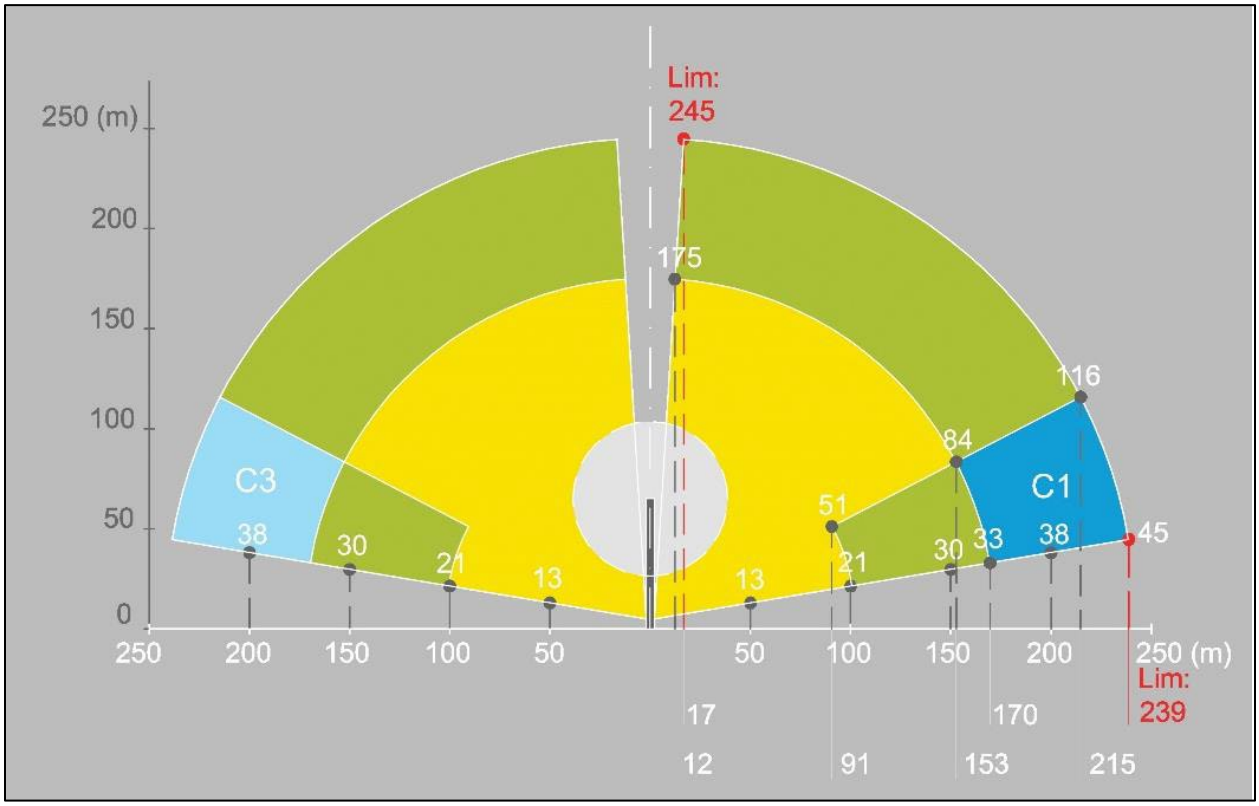

Figure 4. Vertical cross-section illustrating theoretical DTBird detection envelope calibrated for golden eagles, with light gray indicating rotor swept zone, blue indicating detection-only zones, green indicating variable warning-signal trigger zones, and yellow indicating variable dissuasion-signal trigger zones.

Under the DTBird targeting scenario and given calibration for golden eagles, much smaller objects (e.g., small birds and even insects) may trigger detections and deterrents if they are close enough to fill the same number of pixels as a golden eagle would at a much greater distance. Conversely, much larger objects (e.g., airplanes) may trigger detections when they are farther away but fill the requisite number of pixels to be perceived as a possible golden eagle at a relevant distance. Because of these system limitations, false-positive detections and deterrent triggering commonly occur, often at a much greater frequency than events related to target birds (May et al. 2012; Attachment 6).

The DTBird detection and targeting systems incorporate algorithms that reduce false positives caused by factors such as commercial aircraft, insects, and the focal turbine's spinning blades. The constant-pace, arrow-straight flights of high-altitude commercial aircraft are relatively easy to filter out and ignore. Many insects can be filtered out based on their rapid wing beats and erratic flights. Once a specific DTBird installation has been operational for period, a filtering "mask" can be developed that defines the rotor swept area each camera sees and thereby helps the system to filter out false triggers caused by the spinning blades.

Before beginning the Manzana pilot study (H. T. Harvey & Associates 2018), we did not understand that Liquen typically continues adjusting the True False Positive (TFP) filtering algorithms of the DTBird systems they install for as much as an additional 6–8 weeks after they deem the systems fully operational and "commissioned." Although standard practice, once we learned of this additional post-commissioning adjustment practice, we asked Liquen to cease making any further adjustments to create a stable platform for assessing system performance for the remainder of our research at the Manzana site. That point in time was mid-February 2017, approximately 2 months after the Manzana systems were commissioned, which means Liquen had already completed most of the typical post-commissioning adjustments by that time. The false positive performance standard established to guide expansion of this research to the Goodnoe Hills was set based on results derived under this Manzana setup history.

When the Goodnoe Hills systems were setup, we initially requested, once Liquen deemed a given system "fully commissioned", that they make no further algorithm adjustments to establish a consistent and stable platform for our subsequent evaluations. However, a preliminary analysis of the observed false positive rate recorded under this scenario during the first 6.5 months of DTBird operation at the Goodnoe Hills revealed an excessively high rate that greatly exceeded the relevant performance standard for the project. As a result, a proposal was made to the DOE to alter the setup during Year 2 of the overall Goodnoe Hills field study by allowing Liquen to make whatever further adjustments they could to minimize the overall false positive rate. It was agreed that doing so would provide a better basis for comparing DTBird's false positive performance at the two study sites using data collected subsequently at the Goodnoe Hills. Those further adjustments were completed in January 2023.

4.3.3 Sampling Protocol

We present Manzana results based on data collected from January through October 2017 (excludes initial partial month of data from December 2016). We present Goodnoe Hills results based on data collected from September 2021 through July 2023. For each DTBird installation, we randomly selected 10 days per sequential 28-day operational period as our sampling framework. We limited the selections to days when a given turbine and the associated DTBird system were operating at least mostly as expected, with the blades spinning and deterrents triggering when targets were registered to have crossed calibrated distance ranges. For both sites, we excluded from the sample selections all turbine-specific days where and when we conducted UAV flight trials (Attachment 5). On those days, our flight-trial activities undoubtedly influenced the otherwise typical patterns of bird activity around the focal turbines, biasing any other activity observations from those specific days.

4.3.3.1 Classifying Avian Responses to Deterrents

The dataset we developed for this analysis was based on DTBird records that we randomly selected for evaluation to compose a larger experimental analysis (H. T. Harvey & Associates 2018). In 2017, the DAP recorded 19,562 detection events across the seven DTBird installations on days when no UAV flight trials were conducted as part of the pilot study; 8,953 (46%) of these events triggered a deterrent signal. To support investigating the behavioral responses of in situ eagles and other raptors exposed to the deterrent signals, we applied a sampling strategy

to select records to review and classify. Our sampling objective was to amass a temporally and taxonomically representative dataset sufficient to support robust assessments of the probability of deterrence for golden eagles, buteos (mostly red-tailed hawks [*Buteo jamaicensis*] year-round, as well as rough-legged hawks [*B. lagopus*] and ferruginous hawks [B. *regalis*] in Washington and California, respectively, during winter), and all raptors combined.

For each of the functioning DTBird installations, we selected 10 days per 28-day period (the cycling schedule for the larger experiment) across a full year and classifying all detected targets on those days. For evaluating the responses of *in situ* raptors to the deterrent signals, we applied a standardized approach to classifying the responses of all confirmed, suspected, and possible eagles, as well as samples of confirmed turkey vultures (*Cathartes aura*) and buteos for comparison. As described in the previous section, multiple birds occurring simultaneously in the viewsheds of a system's cameras typically confounded rendering precise temporal correlations between detectable changes in the flight behavior of individual birds and the broadcasting of specific warning and dissuasion signals (as reflected in specific triggering timestamps recorded in the DAP). For this reason, we generally excluded event records with multiple birds in view from our deterrence-response classification efforts, as did May et al. (2012). In a few such cases, however, the deterrent signaling could be unambiguously associated with an individual bird of interest, which generally meant the bird was traveling more or less alone and was clearly the only individual that was in a position to trigger the relevant deterrent signal.

To classify deterrence responses, we used the DAP and an on-screen protractor (Straffi 2016) to determine through 2D on-screen measurements whether a bird's flight path appeared to diverge appreciably and away from the RSZ within 5 sec of a warning or dissuasion signal being emitted. For comparative purposes, similar to the approach Liquen personnel typically use to classify deterrence responses, we considered a sustained flight path divergence of >15° away from the deterrent signal that precluded further passage toward the spherical RSZ of the turbine as indicative of a meaningful avoidance response. We also examined the video footage for evidence of correlations between detectable changes in flapping pattern or flight style and emittance of warning and dissuasion signals.

H. T. Harvey & Associates (2018) contains a step-by-step account of the classification process we used to categorize the responses of relevant raptors to the deterrent signals. The process incorporated several subjective and objective criteria for classifying the behavioral response of a given raptor upon exposure to a warning signal and/or dissuasion signal, culminating in a final classification of the response as one of the following:

- *Y: Yes* – reacted in a way that, based on the change in flight pattern and direction, reduced the risk of collision with the turbine blades.
- *P: Potential* – appeared to react to signal, but response was not definitive enough to be confident that the bird was at less risk after signal emission.
- *N: No* – reacted to signal (e.g., temporarily altered its flapping rate) but did not alter its flight path away from RSZ.
- Z: *Not relevant* – did not visibly react to signal.

- U: *Unknown/undetermined* – bird was already moving away from the turbine when the signal was emitted; the video quality or bird image quality was not favorable for determining the 3D reaction of the bird on the 2D video screen; or it simply was not possible to determine with any sense of confidence whether a reaction occurred or not due to other factors.

We excluded from further consideration all cases where we classified the response as "unknown/undetermined."

Along with evaluating behaviors and flight trajectories to classify a bird's response pattern when it triggered a deterrent signal, we classified the potential collision risk the bird was facing prior to triggering a deterrent as follows:

- High – moving toward turbine on a trajectory and at an altitude that could take it near the current RSZ (defined for this purpose as the current, approximate 2D plane of rotation).
- Medium – moving toward turbine on a trajectory and at an altitude that may take it near the turbine, but likely either below or above the RSZ.
- Low – moving perpendicular to or away from the turbine distant from the RSZ, or at high altitude well above the RSZ.

4.3.3.2 Classifying Detected Targets for False Positives Assessment

Once the arrays of turbine-specific sampling days were selected, technicians reviewed the DAP records and videos from those days to classify the targets associated with all detection events recorded while the turbine blades were spinning. Then we focused this multi-site analysis on all such detection events for which the classified target was either a True Fals Positive (TFP) or Nontarget Avian False Positive (NTAFP) that triggered a deterrent signal. False positive detections that do not trigger an audio deterrent may result in excessively cluttered detection databases, which can hamper efficient evaluations of system operation, but they do not run the risk of excessively disturbing nontarget wildlife, wind technicians, and proximate human neighbors or contributing to negative habituation among target species of interest (H. T. Harvey & Associates 2018). Accordingly, for this multi-site assessment we focused exclusively on false positives that triggered deterrent signals.

The technicians classified the targets associated with selected detection events into a broad range of bird species, species groups, and general size categories (species-level identifications were difficult due to low-resolution video records), as well as a range of TFP subcategories. Classification subcategories we lumped together to assess overall TFP detection rates and proportions included several varieties of aircraft (i.e., airplane, helicopter, UAV [excluding our research UAVs], paraglider, and parachute), turbine blades (focal or neighboring turbine), insects, snow, rain, sky artifacts, equipment (i.e., sky artifacts triggered at edges of non-blade turbine features), debris (i.e., floating balloons, paper, plastic bags, etc.), and software/video failures (i.e., poor quality videos preclude target identification). We defined NTAFPs as birds other than large soaring raptors, including abundant common ravens, occasional distinctive falcons (*Falco* spp.) and accipiters (*Accipiter* spp.), and other species ranging from small

passerines to large geese, cranes, and pelicans (plus a few crepuscular bats). Typical large soaring raptors at both study sites were golden eagles, turkey vultures (*Cathartes aura*), red-tailed hawks (*Buteo jamaicensis*), and northern harriers (*Circus hudsonicus*). Less common species at both sites were osprey (*Pandion haliaetus*), Swainson's hawk (*B. swainsoni*; migration and summer only), and ferruginous hawk (*B. regalis*; migration and winter only). Other relevant species unique to each site were abundant rough-legged hawks (*B. lagopus*) and less common bald eagles (*Haliaetus leucocephalus*) during migration/winter at Goodnoe Hills, and rare sitings of California condors (*Gymnogyps californicus*) at Manzana.

The generally poor resolution of the extracted video clips stored in the DAP precluded confidently identifying large proportions of the detected avian targets beyond coarse-scale size/group categories (H. T. Harvey & Associates 2018). Despite intensive QA/QC by the Project Manager/senior avian-raptor expert, nearly 800 Goodnoe Hills records and more than 1200 Manzana records relevant to evaluations of false positives remained classified only as unidentified "big size bird", "unknown medium/large raptor", or "unknown bird", with each classification potentially including some unconfirmed large soaring raptors. To bolster the overall comparative estimates of TFP and NTAFP rates and proportions, we manually classified all unidentified big size birds and unknown birds as either large raptors, medium/large raptors, or NTAFPs based on (a) carefully evaluating representations of other confirmed raptor, raven, and general NTAFP identifications at a given focal turbine on relevant days, (b) considering the general relative abundance of large raptors and ravens at the focal turbine, and (c) making logical assignments based on those considerations. Similarly, we reclassified some records the technicians originally classified as unknown medium/large raptors as large raptors or NTAFPs based on other proximate records identified to species or those two groups.

Partial and complete operational malfunctions of the DTBird systems—caused by several factors—were common at both sites, which led to a variety of sampling imbalances through time and among the different DTBird installations. Operational issues were particularly prevalent at one of the seven Manzana installations (Turbine V17, Figure 1; and see H. T. Harvey & Associates 2018). At the Goodnoe Hills, operational constraints and issues were comparatively rife throughout the study period there. The following constraints were most notable during the 23-month period of record considered in this report:

- System challenges resulted in no useful data being collected at 3 of 14 installations (G29, G51, and G56; see Figure 2) during Year 1 (Attachments 3, 4, and 5).
- The installation at turbine G56 was not fully commissioned until the second *28d Cycle* of Year 2.
- The installation at turbine G48 failed and remained inoperable from mid-November 2022 through early March 2023.
- No useful data were collected at turbine G59 from December 2022 through early April 2023 and at turbine G64 during the month of December 2022.
- The Bonneville Power Administration shut off power to the entire facility from May 1–24, and most of the DTBird systems were not successfully rendered fully operational again until June 6, 2023.
- The installation at turbine G51 was nonfunctional after early July 2023.

- The installation at turbine G67 was largely nonfunctional from early June through early July 2023.
- Most of the installations were largely nonfunctional during the latter half of July 2023.

Given the scale of operational challenges at the Goodnoe Hills, in particular, and the fact that we were not specifically interested in evaluating variation among individual turbines for the assessments herein, we included in our analyses all available and useful data from selected sampling days that met the necessary turbine-DTBird operational criteria for inclusion, as described above. Then we standardized the dependent variables for analysis as the daily counts of TFPs and NTAFPs at each turbine on selected sampling days with relevant records (see Appendixes A and B for summaries of the records used for analysis), and we included *Turbine ID* as a random effect in the statistical models we developed for analyzing variability among the sites and through time. This approach and the robustness of modern analytical models to sampling imbalances and modest violations of distributional assumptions (Schielzeth et al. 2020) helped to reduce potential biases caused by unequal sampling among the sites and DTBird installations.

4.3.4 Controlled Experimental Design at Goodnoe Hills

Data collection began on 1 September 2021 and was expected to continue for two annual rounds of 13 28-day sampling cycles. In the end, sampling was continued for one additional 28-cycle to account for the Bonneville Power Administration having unexpectedly shut down all power to the wind facility from 1–24 May 2023.

The experimental design involved, on a given day, having roughly half of the operational DTBird systems operating in control mode with the deterrent signals not actually broadcasting, and half operating in treatment mode with the deterrent signals broadcasting normally. Here it is important to note that the DTBird systems can be set to trigger and record the timing of deterrent signaling events virtually without the audio deterrents actually broadcasting. Assignments to the control and treatment groups were re-randomized on a daily basis, stratified to ensure (a) daily representation in both the eastern and western halves of the facility, and (b) that each system was operated in treatment mode for at least 10 days per 28-day cycle. Based on preselected rotation schedules (see Attachment 4: Appendix A), Liquen staff implemented and managed automated programming from Spain to control the daily deterrent settings, with necessary daily switching able to occur conveniently during daytime in Spain but nighttime in Washington (DTBird operates only during daylight hours). By randomly assigning treatments on a daily basis and using daily event metrics as the analytical data, we sought to: (1) minimize the potential for turbine-specific habituation; (2) ensure reasonable precision in matching environmental covariate values to response records on a daily basis, rather than seeking to apply covariate values that are averaged or classified across extended periods; and (3) enable effective subsampling of the DTBird event response data.

To select days from which we derived samples used in the analyses, for each operational DTBird turbine we randomly selected 10 days per 28-day cycle for screening, always seeking to the degree possible that each turbine-specific 28-day sample included data for 5 days when the deterrent signals were operating in treatment mode and 5 days when they were operating in

control mode. However, frequent operational failures greatly hindered achieving this intended sampling design. To reduce the effects of frequent system failures in producing unbalanced sampling relative to control-treatment modes, we often adjusted the selected sampling days compared to the initial random selections in an effort to maintain both the 10 days per 28-day cycle sampling objective and 50:50 ratios of control-treatment samples per turbine. Despite these efforts and due to issues beyond our control, the resulting sampling was far from ideal. Nevertheless, especially in this case with Turbine ID treated as a random variable, GLMMs tend to be fairly robust to sampling imbalances as long as the overall representation of data within predictors and covariate classes of interest is relatively robust.

Figure 5. Images portraying the five UAVs deployed during flight trials conducted during this study in California (images A and B) and in Washington (images C–E).

4.3.5 UAV Flight Trials to Assess False Negative Detection Rates

We conducted UAV flight trials at the Manzana site at all seven DTBird installations, with sessions spanning January through August in 2017, but most concentrated in August. We flew flight trials at three Goodnoe Hills DTBird turbines in August 2021 and at four turbines in July 2022. We flew two UAVs during the Manzana flight trials and three different UAVs during the Goodnoe Hills flight trials (see Figure 5). All five UAVs were similar in being fixed-wing plastic/foam-bodied models, with a wingspan, body length, and mass similar to a golden eagle, and painted brown to mimic golden eagle coloration. However, they differed somewhat in overall size, body morphology, and shade of coloration. The Manzana study results suggested that the distance at which the DTBird systems detected the two UAVs flown during those sessions differed significantly, which we interpreted as potentially mimicking differences that could pertain to detecting larger, darker female eagles versus smaller, lighter-colored male eagles (H. T. Harvey & Associates 2018). Accordingly, we purposefully sought to also fly more than one model during the Goodnoe Hills flight trials to support further investigation of this detectability factor. That said, some of the variability in models used stemmed from crashes destroying one of the two aircraft used during the Manzana study and two of the three aircraft used during the Goodnoe Hills study. Further contributing to the variability in UAV models used at each site, the second UAV used during the Manzana study was not available for use during the Goodnoe Hills study.

During both the Manzana and Goodnoe Hills flight-trial efforts, complicated flight conditions for flying light-bodied fixed-wing UAVs and unexpected calamities impinged on our ability to conduct robust suites of UAV flight trials repeated across different seasons with variable sky cover and flight conditions. In the end, both efforts commonly involved concentrated sampling during mid-summer, but differed in that other sampling occurred at the Manzana site at scattered times from mid-January to early March. The extent of sampling across daylight hours also varied at the two project sites. Most flight trial sessions occurred during morning hours when the wind conditions tended to be most compatible for flying fixed-wing UAVs; however, minimal winds allowed for extending the final 2022 sessions at the Goodnoe Hills later into early afternoon (at which point excessive heat precluded further flying for the day).

The key commonality at the two study sites was that we flew primarily pre-delineated linear transects orchestrated as automated flight missions at strategically selected DTBird-equipped turbines, with the goal of achieving representative sampling of the hemispheric, 240-m radius expected maximum-detection-distance envelopes around the sampled DTBird installations. The commonly applied randomized transect selection algorithm delineated flight transects based on multi-layer stratification by compass direction of the flight, flight trajectory (between a maximum 15° ascent and maximum 15° descent), lateral distance from the turbine, and altitude relative to the expected DTBird camera locations. We then packaged collections of 10–20 pre-delineated, turbine-specific transects to orchestrate efficient, single, battery-powered, mostly automated UAV flight sessions using professional pilots, Mission Planner software (ArduPilot Dev Team 2021) on a laptop, and automated radio communication to direct the UAV. Operating several such missions over a multi-hour period composed an individual flight-trial session at a specific turbine, and at both sites we sought to conduct at least half-day flight trial sessions at

several representative DTBird-equipped turbines with compatible landscape settings (i.e., relatively safe places from which to launch and land the UAV, limited topographic complexity, and minimal complications caused by elevated obstacles other than the focal turbine and usually one other adjacent turbine).

Each pre-delineated transect began and ended 100-m line-of-sight distance beyond the projected 240-m detection envelope to support the possibility of detections beyond the expected maximum range. Once the DTBird system targets an object and creates a new detection record in the DAP, no new detection record is created until no additional targeting has occurred for at least 26 seconds. Accordingly, to generate independent transect samples for evaluating the probability of detection and the DTBird system's response characteristics, the automated flight sessions included 30-second loiter periods between each delineated transect at 5–6 preselected, safe destinations located 500 m from the relevant study turbine (previously illustrated in H. T. Harvey & Associates 2018).

Each UAV was equipped with avionics that recorded myriad GPS position, ground and air speed, flight trajectory, and other flight metrics many times per second with high spatiotemporal accuracy. These data were automatically transmitted during the flights to a laptop used to control the automated missions, and could also be extracted directly from the avionics units post-flight. The resulting output from each individual flight was a continuous stream of non-parsed data that had to be translated to a useable format. To extract these data and prepare them for analysis, we followed the detailed procedures and protocols described in H. T. Harvey & Associates (2018). Concisely summarized, this process involved the following primary steps:

1) Translate UAV telemetry log files to spreadsheet format using a publicly available custom program (Fernie 2012).
2) Filter and translate variables recorded by the UAV avionics into useful formats and units of measure, with meaningful variable names.
3) Filter tracking records to:
    a. Exclude data from periods when the UAV was not actually flying (pre-launch and post-landing) or was flying below or loitering outside of detection range.
    b. Include only one record per second to match the resolution of the DAP records.
4) Use ArcGIS 3D Analyst (ESRI, Redlands, CA) to:
    a. Exclude as outliers all loiter-point locations and any other locations recorded at a line-of-sight distance exceeding 340 m; i.e., more than 100 m beyond the expected DTBird maximum detection distance for golden eagles of 240 m.
    b. Code all tracking locations with individual transect numbers based on relevant temporal breaks in the streams of tracking data.
    c. Add additional GIS-derived position metrics and environmental covariates used in analyses.
5) Use the DAP to identify relevant UAV detection and deterrent-triggering event records, and to classify the sky backdrop behind the UAV at the time of each event.
6) Match DTBird detection and deterrent-triggering event records recorded in the DAP to the UAV tracking records based on matching 1-second-resolution timestamps.

7) Finalize datasets for analysis by eliminating all tracking records that are not matched with a DAP event record.

4.4 Analytical Methods

4.4.1 Factors Influencing DTBird Detection Capabilities (Objective 1)

4.4.1.1 Factors Influencing Probability of Detection

To generate estimates of the probability of detecting an eagle-like UAV, we matched DAP detection event records in space and time (resolved to 1-second resolution) with the UAV tracking records to classify each independent UAV flight transect as Detected or Not Detected by the relevant DTBird system. We then calculated the proportions of flight transects detected and not detected at each turbine where we conducted flight trials. The grand-average of the proportions detected then represented the overall estimate of the probability of detecting an eagle-like UAV that passed within the expected 240-meter maximum detection range of the calibrated DTBird systems at each study site, and the converse represented the false negative rate (i.e., the percentage of flights that passed within detection range but were not detected by the DTBird systems).

To generate insight about patterns of variability in the probability of detection, we used ArcGIS tools to calculate the horizontal direction, vertical viewing angle, and line-of-sight (LoS) distance from the detection camera to each individual GPS point along a given UAV flight path, and we used circular statistics to calculate the average *Exposure Direction* (horizontal direction) for each flight transect (Zar 1998). Then we conducted a logistic regression analysis (Systat 13.2.01; Systat Software, Inc., San Jose, CA) with Detected or Not Detected as the binary response variable and several potential predictors considered in the models for evaluation. The relevant predictors were:

- *Site* (Manzana or Goodnoe Hills).
- *Hour of the Day* (e.g., 0900 or 1500 H Pacific Standard Time, using the majority value if the flight segment overlapped two hourly periods).
- *Detection Angle* (°; average vertical angle from camera to UAV).
- *LoS Distance* (minimum line-of-sight distance from camera to UAV).
- *Exposure Direction* (average horizontal angle from turbine to position of UAV, transformed to two orthogonal vectors: sine(*Exposure Direction*) representing a west [negatives values] to east [positive values] vector and cosine(*Exposure Direction*) representing a south [negatives values] to north [positive values] vector).

Given expectations of non-linear relationships from prior site-specific analyses, we considered second-order polynomial terms in the models for *Hour of the Day* and *Detection Angle*, and third-order polynomial terms for *LoS Distance*. We used Akaike's Information Criterion (AIC) scores, individual parameter tests, log-likelihood ratio chi-square tests, and Nagelkerke pseudo-R^2 values to identify the top predictive model given the predictors considered and evaluate the relative influences of various predictors on the probabilities of detection. The logistic GLMMs resulted in predictions of the ln(odds of a response). We used a standard formula

(100*exp[ln[odds]]/[1+exp[ln[odds]]]) to transform the log-odds estimates to probabilities of response (0 to 1 translated to percentages) for the purpose of describing and graphically displaying relationships (Hosmer and Lemeshow 1989).

We also note here that Nagelkerke pseudo-R^2 values do not correlate with typical coefficients of determination R^2 values for non-GLMM models reflecting the proportion of explained variance. Instead, although not well documented in published literature, a typical rule of thumb for interpreting Nagelkerke pseudo-R^2 values is that values ≤2 indicate a weak relationship, values between 0.2 and 0.4 indicate a moderate relationship, and values ≥4 indicate a strong relationship (Shah 2023).

4.4.1.2 Factors Influencing DTBird Detection and Deterrent-Triggering Response Distances

Development of candidate model sets should be guided as much as possible by a thorough understanding of the system being studied (Burnham and Anderson 2010). The multi-site analysis presented here benefited from insights gained from prior site-specific analyses conducted using data collected at the two study facilities (H. T. Harvey & Associates 2018).

The response variable for the analysis was the line-of-sight distance (LoS Response Distance) between the UAV and closest DTBird camera at the time a detection or deterrence event occurred. The operative assumption was that greater response distances can be interpreted as reflecting an improved detection or triggering response, in that earlier (more distant) detection and targeting is expected to provide more time for the deterrents to alter a target bird's behavior well before the risk of collision is acute. We calculated the distances based on the UAV GPS coordinates at the time of the event, using measuring tools in ArcGIS 3D Analyst. Flight samples included in these analyses were necessarily limited to those that triggered a relevant DTBird response. To fit the response-distance data, we built GLMMs and evaluated the influence of various potential random- and fixed-effect predictors. We implemented the models using the 'lme4' package in R (R Core Team 2023; function *lmer,* Bates et al. 2015), with a Gaussian distribution and an identity link function. The initial full model for this analysis had the following structure (see Appendix A for descriptions of each variable):

> LoS Response Distance ~ (1 | *Site* : *Turbine ID*) + (1 | *Site* : *UAV Model*) + *Site* + *Event Type* + *Sky Backdrop* + sin(*Direction from Turbine [DFT]*) + cos(*DFT*) + sin(*Course Over Ground [COG]*) + cos(*COG*) + *Ground Speed* + *Climb Rate* + *Roll Angle* + *Pitch Angle* + *Wind Speed* + *Solar Irradiation* + *Solar Irradiation*^{2} + *Sun Azimuth* + *Sun Elevation* + *Roll Angle* * *Pitch Angle* + sin(*DFT*) * *Sun Azimuth* + cos(*DFT*) * *Sun Azimuth* + sin(*COG*) * *Sun Azimuth* + cos(*COG*) * *Sun Azimuth* + sin(*DFT*) * cos(*DFT*) * *Sun Azimuth* + sin(*COG*) * cos(*COG*) * *Sun Azimuth*

Because the predictor variables were on different scales, we centered and scaled all continuous predictors after applying the following transformations. We transformed *Roll Angles* and *Pitch Angles* to absolute values, expecting that rolling left versus right and pitching up versus down would modify exposure of the UAV profile to the camera similarly. We transformed the *DFT* and *COG* metrics to orthogonal east-west (cos[x]) and north-south (sin[x]) vectors to support linear analyses of these circular variables (Fisher 1995, Cremers and Klugkist 2018). In contrast, we did not similarly transform *Sun Azimuth*, because the range of that variable was only slight

greater than 180° (east in the morning, south at midday, and west in the evening) and therefore did not represent a potential for convergence errors caused by 0° and 360° being equivalent values.

We evaluated *Turbine ID* nested within *Site* (*Site* : *Turbine ID*) and *UAV Model* nested within *Site* (*Site* : *UAV Model*) as random effects, because we expected that DTBird's responses could vary depending on the unique setting at each turbine and variation among the UAVs used, yet neither component was similarly represented at the two sites. In addition, modeling these two factors as random rather than fixed effects acknowledged that the study involved repeated measures (flight sessions) at individual turbines and using different UAVs, such that there was a high likelihood of non-independence among the response distances measured within groupings of these factors. We also modeled *Site* as a fixed effect to determine if DTBird's overall response-distance performance appeared to vary significantly between the two study areas.

We evaluated two- and three-way interactions among the *DFT* and *COG* orthogonal vectors and *Sun Azimuth*, expecting that the influence on response distances of UAV travel direction and directional position from the turbine could markedly depend on the relative position of the sun due to illumination and glare. We also evaluated the two-way interaction between the two UAV "stability" metrics (*Pitch Angle* and *Roll Angle*), anticipating that modeling the interaction of these variables could more accurately reflect the collective influences on exposure of the UAV profile to the cameras than modeling any one metric alone, in part because preventing aircraft stalling effectively precludes maximizing more than one of these variables at the same time. We did not consider any other interactions due to inapplicability and limitations of the available dataset.

To investigate the validity of applying this full model to the multi-site dataset, after we fit the model we used diagnostic tests to evaluate whether the model violated any GLMM assumptions (Zuur et al. 2009, Wood 2017). Specific diagnostics included plotting model residuals to assess independence, equal variances, normal distributions, over- or under-dispersion, and outliers with high leverage. We conducted residual diagnostics using package 'DHARMa' (functions *simulateResiduals*, *plotResiduals*, *testUniformity*, *testDispersion*, *testOutliers*; Hartig 2021). Along with the residual diagnostics, we evaluated potential combinations of predictors for indications of collinearity, and specifically avoided variable combinations that produced variance inflation factors (VIFs) greater than 5 (Hair et al. 1998, Zuur et al. 2010).

To determine the best model for the analysis, we identified the subset of predictors that best explained variation in the observed response distances via stepwise model selection using the step function in R's base 'stats' package (R Core Team 2023) and following the GLMM model selection guidance of Zuur et al. (2009). This stepwise-selection was done in combination with the following criteria to select the best model: ANOVA-based comparisons of nested candidate models, R^2 values, and residual plots. To select final models using Akaike's Information Criterion (AIC), we evaluated only models that met the assumptions of GLMMs. Given the considerable number of predictors and unbalanced categorical factors with some groups having relatively small sample sizes, we used AIC corrected for small sample sizes (AICc) to

compare candidate models to avoid overfitting. We generated graphics resulting from the best model using 'siPlot' Lüdecke 2023) and 'emmeans' (Length 2023), both of which rely on 'ggplot2' (Wickham 2016).

In discussing the significance of statistical results, we label results with $P \leq 0.001$ as *highly significant*, $P \leq 0.05$ as *significant*, and $P \leq 0.10$ as *marginally significant*.

4.4.2 *In Situ* Behavioral Responses of Eagles and Raptors (Objective 2)

Implementing an analogous control-treatment design for evaluating responses to the deterrents was not feasible during the Manzana pilot study. Accordingly, to prepare this multi-site assessment we sought to achieve the following objectives:

A. Use chi-square contingency table analyses with *Site* and categorical *Response* classifications as factors to determine if the apparent responses of eagles and other large raptors to DTBird deterrent signals broadcasted in association with spinning turbine blades differed at the two wind facilities.

B. If the probability of effective deterrence in response to the combination of spinning turbine blades and broadcasted deterrents differs significantly at the two facilities:

    a. Conduct additional logistic generalized linear model (LGLM) analyses to evaluate how various potential predictors influence the probability of effective deterrence at the two sites, limited to the "treatment" data collected at both facilities (i.e., responses to spinning turbines with the deterrents broadcasting).

    b. Conduct no statistical analyses including the "control" data from the Goodnoe Hills site (i.e., responses to spinning turbines with the deterrents muted).

C. If the probability of effective deterrence in response to the combination of spinning turbine blades and broadcasted deterrents does not differ significantly at the two facilities, expand the chi-square and LGLM analyses to include the full combination of treatment data from both sites and control data from the Goodnoe Hills, ignoring Site but including Treatment Group as a predictor. The objective here would be to enhance the single-site control-treatment analysis presented in Attachment 4 by substantially bolstering the available sample size of cases in the treatment group.

D. Develop estimates of the probability of effective deterrence at the two sites that include consideration of the added benefit the DTBird audio deterrents appear to provide above and beyond the effect of spinning turbines alone. The derivation of such estimates will vary depending on whether option (2) or (3) above proves appropriate to pursue.

4.5.2.1 Evaluating Differences in Behavioral Responses Between Sites

To evaluate differences in the categorical responses of raptors to broadcasted deterrent signals at the two study sites, we used 2-way Pearson chi-square analyses performed using the base R package version 4.3.1 (R Core Team 2023). For these analyses, classifications by *Site* (two groups) and *Response* (three groups) categories composed the 2 x 3 contingency tables of interest. If given at least a marginally significant ($P \leq 0.10$) overall chi-square test, we proceeded

to conduct post-hoc comparisons to further characterize the specific *Response* categories within which notable *Site*-specific differences were apparent. For these tests, we used the second post-hoc comparison approach outlined in McDonald (2014). To evaluate the individual significance of the three contrasts of interest, we compared the resulting *P* values to Bonferroni-adjusted values of 0.017 for significance at the overall level of $P \leq 0.05$ and 0.033 for marginal significance at the overall level of $0.05 < P \leq 0.10$.

We prepared these chi-square analyses for all analyzed cases, all confirmed/probable golden eagles, all confirmed/probable turkey vultures, and all confirmed/probable buteos. Further, the datasets included three possible response variables, one pertaining to responses to warning signals alone, one pertaining to responses to dissuasion signals alone, and one including responses to single deterrents or to the combination of both deterrents signaling in sequence, where applicable. For this multi-site analysis we focused only on the combined response data to maximize sample sizes and emphasize the overall effects of the deterrent system. In a few cases, the resulting cell sample sizes were small, but Pearson chi-square tests are known to be robust as long as expected cell frequencies exceed 1.0 (Jeffreys 1939), and our preliminary investigations showed no notable differences in outcome using the alternative Fisher's Exact Test. We did not strive to develop more complicated 3-way chi-square statistical models that included consideration of relative collision risk prior to deterrent triggering as a third predictor (H. T. Harvey & Associates 2018). However, we ultimately addressed this important potential influence again using a LGLM approach.

4.5.2.2 Evaluating Factors Influencing Behavioral Responses to Deterrents

As described further below, the initial chi-square analyses indicated that the probability of effective responses to broadcasted deterrents was often lower at the Goodnoe Hills facility than at the Manzana facility. Therefore, pursuing the second phase of Objective B rather than Objective C, as outlined above, was warranted. Accordingly, we did not seek to integrate the treatment data from both sites to compare against the control data generated only at the Goodnoe Hills. Instead, we sought to develop further insight about possible drivers of the difference in the probability of effective responses to broadcasted deterrents at the two sites by composing LGLM analyses to evaluate the influences of several potential predictors. These analyses were necessarily limited to cases involving responses to broadcasted deterrents. Further, we collapsed the *Response* variable from four to two categories to compose a binary response variable for the LGLM analysis: 1 = probable effective response (*CE* + *PE* classifications as described above) and 0 = no effective response (*I* = *N* + *Z* classifications). We prepared two analyses—one based on the multi-species dataset and one limited to probable golden eagles—and focused only on the combined deterrence response classifications. For the multi-species analysis, we included a *Species Group* variable in the model to highlight potential differences among the three primary species groups: eagles, vultures, and buteos. To facilitate evaluation of *Species Group* as a predictor, we reduced the dataset to only those cases that we could confidently identify as belonging to one of these three groups. The initial full model for the multi-species analysis was as follows:

> ln(Odds of effective deterrence) ~ *Site* (Manzana CA or Goodnoe WA) + *Species Group* (Eagle, Vulture, or Buteo) + *Preexposure Risk* (risk of exposure to turbine before

deterrence: low, medium, or high) + *Wind Speed* (meters/second; measured by turbine anemometer) + all possible 2-way interactions

The initial full model for golden eagles was the same except for excluding the *Species Group* variable. We implemented the LGLM analyses using the 'glm' function in R (R Core Team 2023). To settle on final models, we used likelihood ratio tests for individual parameters and compared Akaike Information Criterion (AIC) scores for all possible candidate models reflected in the full model statements to identify the most parsimonious combinations of predictors (Burnham and Anderson 2010). In considering the merits of different candidate models, we also used diagnostic residual plots to evaluate conformity to the assumptions of LGLMs, plots of model residuals versus leverage and Cook's distance to identify potential outliers, and McFadden's pseudo-R^2 to assess the explanatory power of models (McFadden 1974, Friendly and Meyer 2016).

The LGLM resulted in predictions of the ln(odds of effective deterrence). We used a standard formula (100*exp[ln[odds]]/[1+exp[ln[odds]]]) to transform the log-odds estimates to probabilities of response (0 to 1 translated to percentages) for the purpose of describing and graphically displaying relationships (Hosmer and Lemeshow 1989).

4.4.3 Factors Influencing False Positive Detection Rates (Objective 3)

We used R 4.3.2 (R Core Development Team, Vienna, Austria) to develop generalized linear mixed models (GLMMs) illustrating variation in TFP and NTAFP rates at the two study sites. We developed independent analyses for TFPs and NTAFPs, focusing on four model constructs for TFPs and two model constructs for NTAFPs. Given the additional false-positive filtering adjustments made during Year 2 of the Goodnoe Hills study, our first analytical objective was to compare TFP and NTAFP rates at the Goodnoe Hills across comparable periods of Year 1 and Year 2. Then we analyzed differences between the two study sites by comparing results from the Manzana site against results from only a comparable period of Year 2 at the Goodnoe Hills. For both sets of comparisons, we analyzed two models with the following variable structures:

Goodnoe Hills Year 1 versus Year 2

TFPs/Turbine/Day ~ (1|*Turbine ID*) + (1|*28d Cycle:Date*) + *Year* + *28d Cycle* + *Year*28d Cycle*

NTAFPs/Turbine/Day ~ (1|*Turbine ID*) + (1|*Month:Date*) + *Year* + *Month* + *Year *Month*

Manzana versus Goodnoe Hills Year 2

TFPs/Turbine/Day ~ (1|*Turbine ID*) + (1|*28d Cycle:Date*) + *Site* + *28d Cycle* + *Site*28d Cycle*

NTAFPs/Turbine/Day ~ (1|*Turbine ID*) + (1|*Month:Date*) + *Site* + *Month* + *Site*Month*

We included *Turbine ID* as a random effect in all models to account for uncontrolled variation resulting from the unique spatial and temporal influences of individual turbine locations and to avoid pseudo replication, and we treated *Date* as a random categorical factor nested within *28d Cycle* or *Month* to account for the influence of variable sampling days and avoid pseudo replication. We examined the models with *28d Cycle* and *Month* as alternative temporal

predictors to address different interests in examining patterns of variation though time. Specifically, we used *28d Cycle* to evaluate the influences of operational duration on TFP rates, and we used *Month* to evaluate the seasonal influences of specific times of year on the prevalence of NTAFPs (including natural factors that vary seasonal, such as precipitation and insects).

We analyzed these data using negative binomial GLMMs, which account for typical overdispersion of count-based data. We used the 'glmmTMB' package in R (Brooks et al. 2017a, b; Magnussen et al. 2022) to generate the models with a log-link. The negative binomial response distribution ('binom2', with variance = μ[1+μ/k], where μ is the mean and k is the overdispersion parameter) accounted for overdispersion in the data.

We tested for differences in daily counts among *28d Cycles* or *Months* using chi-squared maximum likelihood-ratio tests to evaluate the significance of the fixed factors in the models. To obtain estimated means for daily turbine-specific TFP and NTAFP counts based on the selected final models, we used the 'ggpredict' function ('ggeffects' package; Lüdecke et al. 2022). We identified differences among means using planned post-hoc comparisons following Tukey's Honestly Significant Differences test (Tukey 1949) to maintain a family-wise alpha of 0.05. The planned comparisons were limited to pairwise comparisons among *28d Cycles* or *Months* within *Years* or *Sites*.

4.4.4 *In Situ* Experimental Evaluation of Raptor Responses (Objective 4)

The research hypotheses we formulated for the experiment were as follows:

Hypothesis A: The probability of an eagle triggering a dissuasion signal will be lower for DTBird turbines operating in treatment mode (deterrent signals broadcasting) compared to those operating in control mode, because broadcasted warning signals deter target raptors from approaching closer and triggering a dissuasion signal.

Hypothesis B: The average dwell time of eagles in the vicinity of DTBird-equipped turbines—as reflected in the length of relevant targeting videos recorded by the DTBird detection system—will be reduced around systems operating in treatment mode compared those operating in control mode, because broadcasted deterrent signals discourage birds from lingering near focal turbines.

Hypothesis C: The probability of an eagle crossing the active rotor swept area (RSA) of DTBird-equipped turbines will be lower for systems operating in treatment mode compared to those operating in control mode, because operation of the deterrent signals reduces the likelihood of target raptors entering the RSZ of turbines.

To analyze the full two-year experiment dataset, we used generalized linear mixed models (GLMMs) to evaluate the three research hypotheses using different response variables: 1) binary logistic response = whether or not a detected large raptor triggered a dissuasion signal, 2) continuous response (seconds) = tracking video length per large raptor targeting event, and 3) binary logistic response = whether or not a detected large raptor appeared to cross through or close to the RSA. Challenges producing a consistent, accurate, and robust dataset on

possible RSA crossings based on interpreting 3D responses from 2D video images limited our ability to evaluate research Hypothesis C.

Our GLMM designs considered DTBird turbines to be sampling units and included *Turbine ID* as a random effect in the models to account for inherent, localized, spatial variation in the landscape settings and eagle/raptor activity patterns at different turbines. All models also included sampling date nested within *Turbine ID* to account for highly variable temporal sampling at each turbine and inherent, localized, temporal variation in the environmental conditions, human activity patterns, and other factors that likely influenced the activity patterns and responses of target raptors around individual turbine locations. For this purpose, we transformed sampling dates to *Elapsed Days* since projection inception.

Given frequent uncertainties in species-specific identifications and attendant sample-size limitations for focal golden eagles, we developed independent models for three hierarchical taxonomic groups to provide effective insight: 1) confirmed and probable golden eagles, 2) confirmed and probable golden and bald eagles, with *Species* considered as a potential predictor; and 3) all confirmed and probable eagles, including unidentified eagles, without considering species as a potential predictor.

Predictors and covariates considered in the GLMMs were as follows:

- Random effects:
    - *Turbine ID*.
    - *Days Elapsed nested within Turbine ID*.
- Fixed effects:
    - *Treatment Group* (binary): treatment or control.
    - *Species* (categorical): included in models focused on confirmed golden and bald eagles combined, but excluded from models focused on golden eagles alone and all possible eagles, including those not confirmed to species.
    - *28-day Cycle* (discrete continuous): sequential series from 1 to 27 over 25-month period, with period 23 mostly not represented due to an unanticipated 1-month facility shut down.
    - *Time of Day* (continuous, Pacific Standard Time, translated to minutes of the day): second order term included to account for expected curvilinear relationship.
    - *Cloud Cover* (categorical): reflecting predominant daily condition gleaned from review of DTBird video records and coarsely classified by technicians as fair (mostly cloud free), partly cloudy (<50% cloud cover), cloudy (≥50% cloud cover with distinctly variable cloud definitions and brightness), or overcast (complete and largely uniform gray or darker cloud cover).
    - *Wind Speed* (continuous, meters/second): derived from turbine system metrics and averaged across duration of tracking event.
    - *FPs per Day* (discrete continuous): number of daily deterrent-trigger events resulting from false positives, including both true false positives (non-bird, including inanimate moving/flying objects, insects, precipitation, and sky artifacts) and non-target avian false positives (non-focal birds).

The selected covariates represented factors that: 1) were discernable using the DTBird DAP or were attainable from the wind facility; 2) we expected to have the potential to influence the ability of focal raptors to visualize the turbines and hear and respond to the deterrents; and 3) could influence the responses of focal raptors by increasing the frequency of deterrents being broadcasted. Given focal interest in evaluating *Treatment Group* as a predictor, we also evaluated all possible two-way interactions between *Treatment Group* and the other potential predictors/covariates. For all continuous independent variables, we centered and scaled the values as (value - mean)/SD prior to analysis.

For each species group, we developed GLMMs to test for the effects of *Treatment Group* and the five potential covariates on the three dependent variables. We used the R function 'glmer' in the lme4 package (Bolker 2023) to compile and evaluate GLMMs based on a binomial error distribution with a logit link (i.e., mixed-effects logistic regression), and maximum likelihood estimation with the *bobyqa* optimizer and the maximum number of function evaluations set to 10^5, to model the probability of detection events triggering a dissuasion signal and whether or not an RSA cross occurred. We used the R Package 'glmmTMB' (Brooks et al. 2023) to compile and evaluate GLMMs based on a gamma error distribution with a log link and maximum likelihood estimation to analyze dwell time (recorded video length) as a dependent variable. We compared Akaike's Information Criterion (AIC) scores for candidate models to balance considerations of model fit and parsimony (considering a ΔAICc of ≤2 points indicative of similarly competitive models) and used Wald *z*-tests and Drop1 likelihood-ratio chi-square tests to further assess the relative importance of different predictor variables and ultimately identify a top model for each independent analysis (Burnham and Anderson 2002, Bolker et al. 2009, Symonds and Moussalli 2011).

To ensure a good model fit, normally distributed residuals, and homogeneous variances, we inspected residual plots for the selected models and individual grouping factors by plotting results using the 'simulateResiduals' function (package 'DHARMa'; Hartig 2019) applied to the selected model. We also conducted goodness-of-fit tests on these residuals using the 'testUniformity' function from the same package, which performs a Kolmogorov-Smirnov test for specified factors and combinations of factors (including the overall model) to evaluate conformity to a normal distribution. We used the functions 'testOutliers', 'testOverdispersion', and 'testZeroInflation' to confirm that the residuals did not include outliers nor exhibit overdispersion or zero-inflation (Hartig 2019).

To evaluate Wald *z* tests and Drop1 likelihood ratio chi-square parameter tests for individual predictors considered during GLMM development, we adopted $P \leq 0.10$ as our threshold for retaining predictors in the selected models. We chose this relatively liberal threshold to ensure representation of potentially noteworthy relationships that might have emerged more strongly had our sampling not suffered from frequent spatial and temporal imbalances in the operation of the study installations and resultant sampling, and uncertainties pertaining to species identifications. We refer to tests and contributions as marginally significant if $0.05 < P \leq 0.10$, significant if $0.01 < P \leq 0.05$, and highly significant if $P \leq 0.01$.

For the logistic GLMMs, which resulted in predictions of the ln(odds of a response), we used a standard formula (100*exp[ln[odds]]/[1+exp[ln[odds]]]) to transform the log-odds estimates to probabilities of response (0 to 1 translated to percentages) for the purpose of describing and graphically displaying relationships (Hosmer and Lemeshow 1989).

4.4.5 Multi-site Analysis of Collision Risk Reduction (Objective 5)

We used data generated by the two-site DTBird evaluations and the controlled experiment at Goodnoe Hills to quantify DTBird's effect on golden eagle collision risk, as described above. We initially intended to translate our results to applying the Bayesian collision risk model (CRM) recommended by the U.S. Fish and Wildlife Service (2013; and see New et al. 2015), using eagle flight times recorded by DTBird at control and treatment turbines as a proxy for eagle activity. However, we found comparisons of proportional responses to be most germane, because any estimates we could generate portraying absolute reductions in the number of eagles killed per year would be site specific, whereas proportional estimates have the potential to be applied across sites based on site-specific fatality projections.

4.4.6 Performance Reliability and Cost Analysis (Objective 6)

A more detailed breakdown of costs to purchase, acquire, install, and maintain DTBird is detailed in Attachment 10.

4.5 Results

4.5.1 Factors Influencing Probability of Detection (Objective 1)

The sample sizes of independent site- and turbine-specific UAV flight transects that formed the basis for quantifying and investigating variation in the probability of detection ranged from 144–221 samples per turbine at the Manzana site and 54–131 samples per turbine at the Goodnoe Hills site (Table 1). At the Manzana site, DTBird detected 798 of 1,279 (62%) UAV flight transects, with the detected proportions ranging from 47–75% across seven sampled turbines. At Goodnoe Hills, DTBird detected 310 of 481 (64%) UAV flight transects, with the detected proportions ranging from 56–80% across five sampled turbines (Table 1).

Table 11. Numbers of UAV flight transects by sampled turbine analyzed to quantify and investigate variation in the probability of DTBird detecting an eagle-like UAV at the Manzana Wind Power Project in California and Goodnoe Hills Wind Farm in Washington.

Site	Turbine	Detected	Not Detected	Total	% Detected
Manzana	D01	80	64	144	56
	D04	129	62	191	68
	D08	106	65	171	62
	E11	143	54	197	73
	T13	116	38	154	75
	U7	130	91	221	59
	V17	94	107	201	47
Subtotal		798	481	1,279	62
Goodnoe Hills	G34	65	16	81	80

	G44	81	50	131	62
	G58	69	36	105	66
	G64	33	21	54	61
	G75	62	48	110	56
Subtotal		310	171	481	64
Total		1,108	652	1,760	63

The final model derived to illustrate the influence of spatial and temporal predictors on the probability of detection based on UAV flight trials had the following form:

$$\ln(\text{Odds of Detection}) \sim \textit{Site} + \textit{Hour of the Day} + \textit{LoS Distance} + \textit{LoS Distance}^2 + \textit{LoS Distance}^3 + \textit{Detection Angle} + \textit{Detection Angle}^2$$

The log-likelihood ratio goodness-of-fit test comparing the selected model and null model indicated a highly significant fit (χ^2 =476.7, df = 7, $P < 0.001$) and the Nagelkerke Psuedo-R^2 for the model was 0.324, indicating a moderate relationship. Comparisons with other candidate models are illustrated in Attachment 6: Appendix C, and coefficients, parameter tests, and diagnostics for the selected model are presented in Attachment 6: Appendix D.

The selected model indicated that the probability of detection:

- Averaged higher at Goodnoe Hills than at Manzana (discussed further below).
- Increased as the day progressed, from an average of approximately 57% during the 06:00 H to 75% during the 20:00 H (Figure 6).
- Was highest (estimated average ~75%) when the LoS Distance to a flight track was 50–75 meters from the cameras; decreased slightly at closer distances; and decreased at greater distances down to an estimated average of approximately 50% at the 240 meter

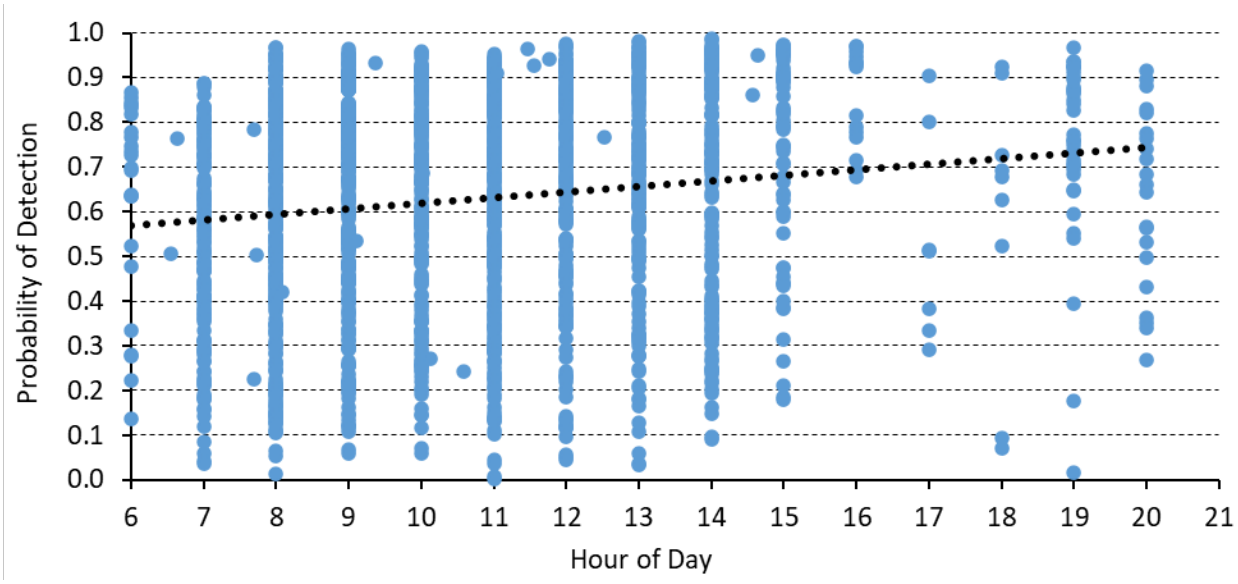

Figure 6. Modeled Linear Relationship Between Predicted DTBird Detection Probabilities for Individual UAV Flight Transects and Hour of the Day.

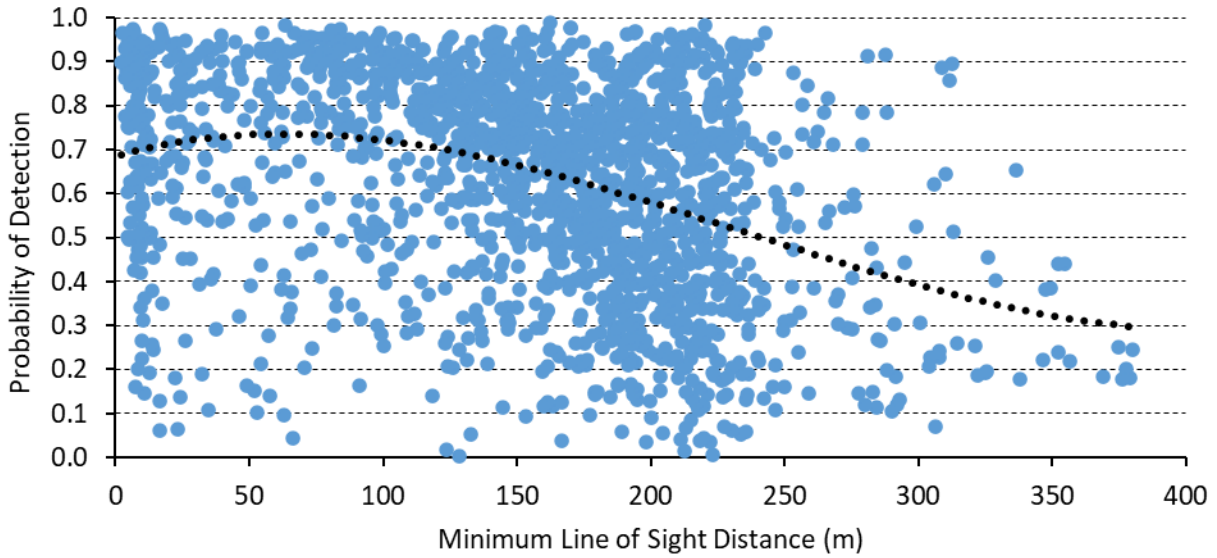

Figure 7. Modeled Third-Order Relationship Between Predicted DTBird Detection Probabilities for Individual UAV Flight Transects and the Minimum Line-of-Sight Distance to the DTBird Camera.

expected (calibrated) maximum detection distance for targets the size of golden eagles, but remained at an estimated 30% as far out as 380 meters from the cameras (Figure 7).

- Was highest (estimated mean ~65%) when the Average Detection Angle from the camera to a flight track was moderate (approximately 20–30° above horizontal from the camera) and decreased on average by 25–35% at minimum lower and maximum higher observed angles (Figure 8).

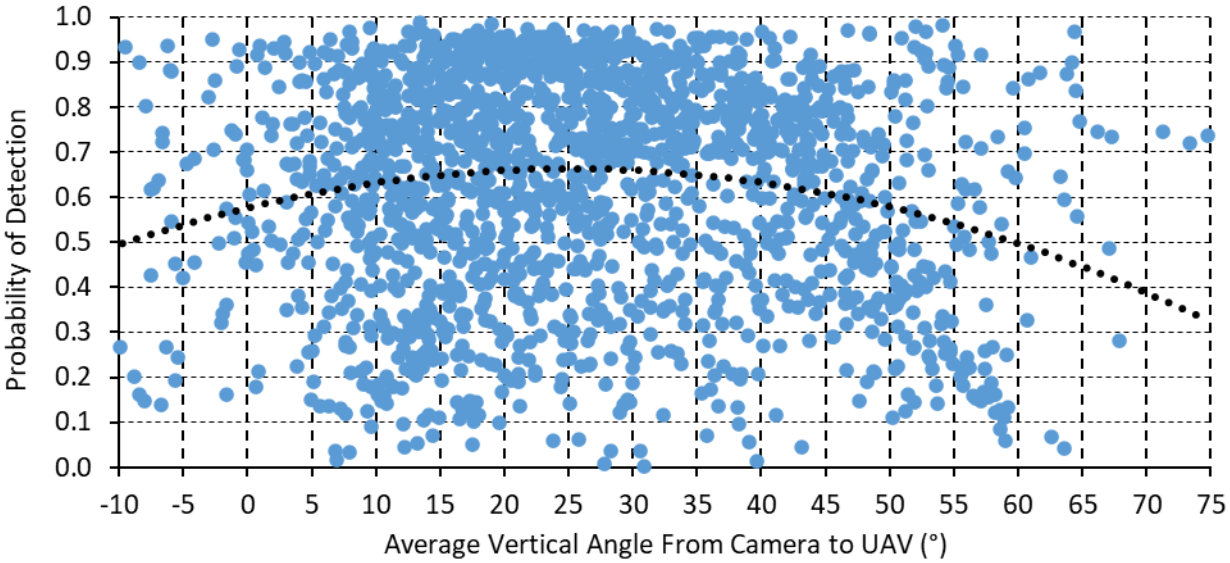

Figure 8. Modeled Second-Order Relationship Between Predicted DTBird Detection Probabilities for Individual UAV Flight Transects and the Average Vertical Angle from the DTBird Camera.

Based on the model output and the range of flights considered in formulating that model, the overall average probability of detecting an eagle-like UAV at the two study sites was 63 ± 1.1% (95% CI). However, as the basis for the predictive model, we included a broad range of flights with LoS Distances extending out as far as 380 meters, including DTBird detection distances of up to 375 meters. The intent was to maximize good model fit by including useful data that extended spatially beyond the focal, calibrated maximum detection distance of 240 meters. For the purpose of comparing the estimated overall probability of detection (or conversely false negatives) against the performance standard established for this project (63%), a fairer metric is the probability of detecting an eagle-like UAV that flies within 240 meters or less of the cameras. Based on model output and this restriction, the relevant detection probabilities were 66 ± 1.3% (95% CI) at Manzana, 64 ± 1.9% at Goodnoe Hills, and 65 ± 1.1% overall. Note that these indicators suggest that the probability of detection was slightly higher at the Manzana site, whereas the modeled full dataset suggested the opposite (Table 1), emphasizing that any difference between the two sites was at best marginal.

Flipped about to focus on false negatives, these results suggest that the probability of DTBird missing a detectable flight was overall <20% when the LoS Distance to the flight was between approximately 30–120 meters, <30% at distances of <20 meters and between 120–160 meters from the cameras, and exceeded 50% only beyond 200 meters. Otherwise, flights were missed more often at the Goodnoe Hills, during morning light, and at both low and high detection angles.

4.4.1.2 Factors Influencing DTBird Detection and Deterrent-Triggering Response Distances

The flight trials conducted at the Manzana study site in 2017 occurred at all seven DTBird installations between 06:45 and 16:45 H Pacific Standard Time (PST) on 2 days in mid-January, 3 days in late February and early March, and 5 days in August (Table 2). The January and February/March flights involved an initial, custom-built aircraft (AES Custom; Figure 5A) flown by our first pilot, but unfortunately that aircraft crashed and was damaged beyond repair during the March flights. The August flights then involved a different pilot and custom-built aircraft (AUV Custom; Figure 5B). The Manzana missions resulted in a total of 1,279 usable, distinct flight segments (Table 2).

Table 22. Summary of UAV Flight Trials Conducted at the Manzana Wind Project Site in California that Contributed Data for Analysis.

Date	Sample Period (PST)	Turbine	Aircraft[1]	Missions Flown	Yield of Transect Samples
17-Jan-2017	08:15–11:40	V17	AES Custom	3	55
	13:05–16:45	E11	AES Custom	4	73
18-Jan-2017	08:45–12:05	D4	AES Custom	4	69
	13:15–14:25[2]	D8	AES Custom	2	32
21-Feb-2017	07:55–12:05	U7	AES Custom	6	94
	13:15–13:50[2]	D1	AES Custom	1	18
28-Feb-2017	10:45–15:45	T13	AES Custom	6	105
01-Mar-2017	08:35–10:10[3]	E11	AES Custom	2	31
07-Aug-2017	07:35–13:55	V17	AUV Custom	8	146

08-Aug-2017	07:05–13:05	D8	AUV Custom	7	139
	13:55–15:50	U7	AUV Custom	2	37
09-Aug-2017	07:05–11:30	D4	AUV Custom	6	122
	12:35–13:15[3]	U7	AUV Custom	1	16
10-Aug-2017	06:45–12:10	D1	AUV Custom	8	126
	13:00–15:00	T13	AUV Custom	3	49
11-Aug-2017	06:35–08:40	U7	AUV Custom	3	74
	09:25–12:25	E11	AUV Custom	5	93
			Totals	71	1,279

[1] See Figure 5 for pictures of the aircraft.

[2] Aborted prematurely because of excessive wind or inclement weather.

[3] Aborted prematurely because of UAV operational failure.

At the Goodnoe Hills study site, the flight trials conducted in 2021 occurred at three turbines on two consecutive days in early August, involved a new pilot and mixed use of two UAVs (Clouds [Figure 5C] and Believer [Figure 5D]), and resulted in 210 flight samples suited to analysis (Table 3). Unfortunately, this flight trial session was terminated prematurely when both aircraft suffered fatal crashes. We also attempted an initial round of flight trials at this site in May 2021, but we were generally unable to proceed due to wind speeds that were incompatible with conducting flight trials with light-bodied UAVs. The flight trials conducted in 2022 then occurred at four turbines on four days in late July. They involved another piloting team and limited use of another Clouds aircraft, but primarily a new Ranger aircraft (Figure 5E), and resulted in 272 flight samples suited to analysis. We also conducted another apparently successful series of eight flights at turbine G51 during the trial session in July 2022, only to find out later that a DTBird hardware mismatch issue resulted in no recordings of those flights. Thus, our sampling at this site fell short of expectations, which we could not overcome due to budget limitations.

Table 33. Summary of UAV flight trials conducted at the Goodnoe Hills Wind Farm study site in Washington that contributed data for analysis.

Date	Sample Period (PST)	Turbine	Aircraft[1]	Missions Flown	Yield of Transect Samples
02-Aug-2021	07:42–08:46	G58	Believer	2	38
	11:05–13:04	G58	Clouds	2	67
	17:43–20:33	G34	Clouds	3	71
03-Aug-2021	08:34–09:29[2]	G44	Believer	2	34
25-Jul-2022	11:57–12:10[2]	G34	Clouds	1	10
26-Jul-2022	09:59–15:55	G64	Ranger	4	54
27-Jul-2022	08:15–15:41	G75	Ranger	7	111
29-Jul-2022	07:49–13:40	G44	Ranger	8	97
			Totals	29	482

[1] See Figure 5 for pictures of the aircraft.

[2] Aborted prematurely because of UAV operational failure.

The evaluation results for the initial full model and other models considered as part of the backward selection process used to identify the best model are portrayed in Appendix B. The final, selected model had the following form:

> LoS Response Distance ~ (1 | *Site* : *Turbine ID*) + (1 | *Site* : *UAV Model*) + *Site* + *Event Type* + *Sky Backdrop* + *Ground Speed* + *Wind Speed* + *Roll Angle* + *Pitch Angle* + *Roll Angle* * *Pitch Angle*

A model with only the random effects included (AICc = 20010.06) reduced the AICc score by a substantial 223.48 points compared to the null model (AICc = 20233.54), and the selected model (AICc = 19918.34) reduced the AICc score by another substantial 91.2 points (315.2 total points compared to the null model). These results confirm noteworthy improvements in balancing parsimony and explanatory power (Burnham and Anderson 2010). The selected model also reduced the AICc score by 70.9 points compared to the full model (AICc = 19989.19), further reflecting a markedly improved model. However, the Nakagawa marginal pseudo-R^2 for the model (0.092) was low (Nakagawa and Shielzeth 2013), indicating that the included fixed effects provided only marginal explanatory power and a lot of variability in the dataset remained unexplained.

Diagnostics indicated that the final model satisfied the important assumptions of independence, normally distributed residuals, and the absence of significant collinearity among the predictors. However, Levene Tests for homogeneity of variances across groups within categorical variables (Zuur et al. 2009, Hartig 2021) confirmed modest deviations from ideal for *Site* and *Event Type*, but not for *Sky Backdrop*. These results suggest that the assumption of homogenous variances within groups was not completely met. Nevertheless, by incorporating random effects in the model, GLMMs estimate the variance components for the random effects, capturing the variability between groups and within groups. This flexibility in modeling allows for the accommodation of heteroscedasticity and helps to mitigate the impact of violations of the assumption of homogeneity of residual variances. Additionally, GLMMs can provide accurate parameter estimates and valid statistical inference even in the presence of heteroscedasticity; the mixed-effects structure helps to account for the correlation structure within the data, which reduces bias and provides robust standard errors for hypothesis testing (Zuur et al. 2009).

Output for the selected model indicated that including *Site* : *Turbine ID* as a random effect accounted for modest variation among turbines in modeled response distances (Figure 9). Specifically, the modeling results suggested that response distances were more variable among the seven Manzana turbines than among the five Goodnoe Hills turbines. Among the seven Manzana turbines, response distances were approximately 8.9 m shorter than the estimated global average at one turbine (V17), 7.7 m longer than average at one turbine (T13), and values for the other five turbines ranged from -0.9 m shorter to 1.7 m longer than the grand average. In comparison, the range of variation among the five Goodnoe Hills turbines was from 4.5 m shorter to 3.9 m longer than average, and values for the other three turbines ranged from -1.1 m shorter to 2.4 m longer than average. Although noteworthy but not particularly substantial

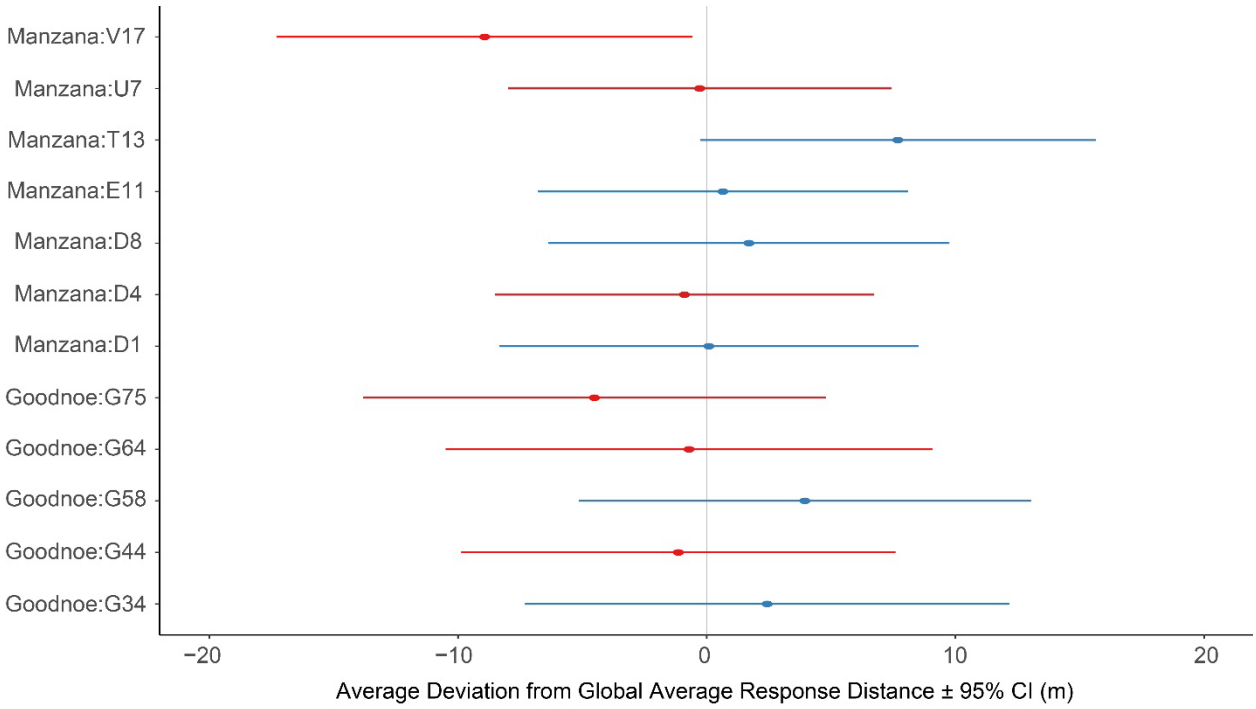

Figure 9. Deviations from the estimated global average DTBird response distance associated with different site-specific turbine installations, estimated as a nested random effect in the multi-site GLMM developed for the study.

differences, these apparent turbine-level variations likely reflect situation-specific landscape variation leading to modest variability in DTBird's ability to detect and target objects of interest.

Output for the selected model indicated that including *Site* : *UAV Model* as a random effect also captured noteworthy variation in the global average response levels attributable to the different UAV models used (Figure 10). The two UAV models used at the Manzana site showed the greatest variance in response distances: approximately 15.0 m shorter than the estimated global average across UAV types for the AUV Custom aircraft (with a skinny tubular hind body and more variable coloration; Figure 5A) and 15.0 m longer than average for the AES Custom aircraft (overall a more eagle-like torso and darker coloration; Figure 5B). At the Goodnoe Hills, variation among the three UAV models was less pronounced, ranging from an estimated 5.1 m shorter than average for the Believer aircraft (a relatively heavy, dark, and fast-flying aircraft; no picture available), 4.2 m longer than average for the Clouds aircraft (a relatively large and robust body and intermediate coloration; Figure 5D), and a nominal 0.9 m longer than average for the Ranger aircraft (longest wing span, but relatively narrow features and intermediate coloration; Figure 5C).

The coefficients and associated parameter tests for the fixed effects retained in the selected model are provided in Table 4. The selected model suggested that the retained fixed-effect predictors influenced the DTBird LoS Response Distances as summarized below.

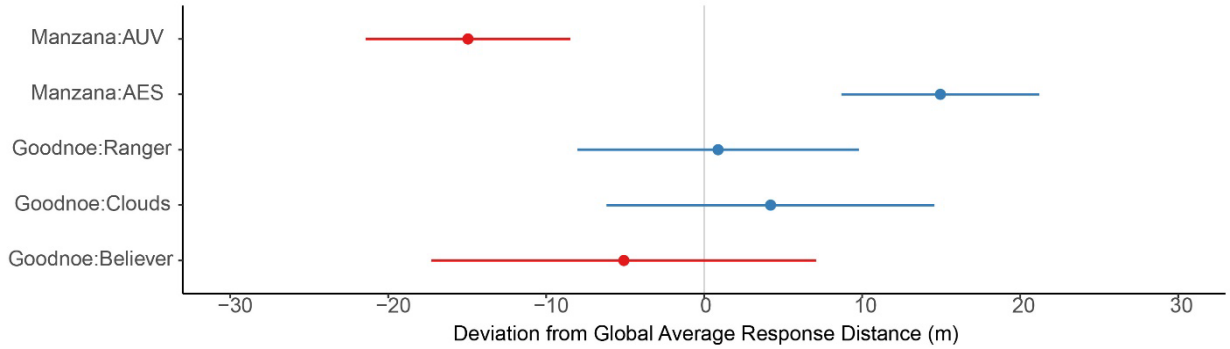

Figure 10. Deviation from the estimated global average DTBird response distance associated with site-specific use of different UAV models, estimated as a nested random effect in the multi-site GLMM developed for the study.

Table 44. Coefficients and parameter t-test results for fixed effects represented in the selected multi-site GLMM with DTBird response distance as the dependent variable.

Predictor	Coefficient	SE	df	*t*	*P*
(Intercept)	197.677	9.312	5.0	21.2	<0.0001
Site : Manzana [1]	-32.701	13.621	3.7	-2.4	0.0794
Event Type : Warning [2]	0.755	4.314	1798.7	0.2	0.8612
Event Type : Dissuasion [2]	-14.149	3.412	1793.9	-4.1	<0.0001
Sky Backdrop : PartlyCloudy [3]	3.900	5.751	48.9	0.7	0.5008
Sky Backdrop : MostlyCloudy [3]	10.864	5.980	104.6	1.8	0.0721
Sky Backdrop : Overcast [3]	19.361	5.433	105.1	3.6	0.0006
Ground Speed	3.282	1.595	1744.8	2.1	0.0397
Wind Speed	3.229	1.657	1623.0	1.9	0.0515
Roll Angle	2.459	1.418	1798.4	1.7	0.0830
Pitch Angle	-0.719	1.429	1800.1	-0.5	0.6148
Roll Angle * Pitch Angle	-5.607	1.315	1796.0	-4.3	<0.0001

[1] Reference category: Goodnoe Hills.
[2] Reference category: Detection event.
[3] Reference category: Fair skies.

***Site*:** The coefficient and parameter test for this fixed effect suggested that response distances averaged marginally shorter overall at the Manzana site than at the Goodnoe Hills site, and the post-hoc comparison of estimated means and variances illustrated that difference, but confirmed that it was not significant at $P \leq 0.05$ (Figure 11).

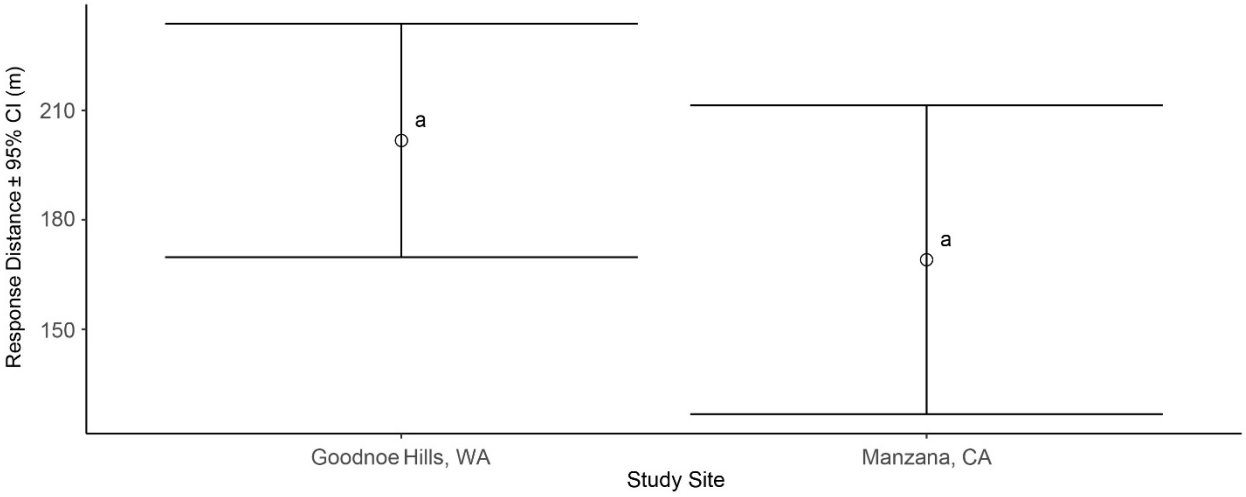

Figure 11. Modeled relationship between DTBird response distances and study site, with shared letters indicating pairwise differences that are not significant at *P* ≤ 0.05.

***Event Type*:** Including *Event Type* as a fixed effect accounted for the significant "structural" (i.e., a system calibration/programming feature) difference in expected trigger distances for dissuasion signals compared to initial detections and warning signals (Figure 12). Calibrated for this study, initial detections were expected to occur at 240 m from the cameras throughout the projected detection envelope, while warning signals were also to be triggered at 240 m throughout the core envelope and at 170 m across lower, outer reaches of the detection envelope (see Figure 4 and H. T. Harvey & Associates 2018 for graphical illustrations). In contrast, dissuasion signals were expected to trigger at 170 m from the cameras throughout most of the expected detection envelope, and at 100 m across lower, outer reaches of the detection envelope. In contrast, the marginal means produced from the model for this parameter reflected the difference in average response distances for dissuasion signals (175.7 ± 7.34 m [SE]) and the comparatively minimal difference between the average response distances for initial detections (189.9 ± 7.00 m) and warning signals (190.61 ± 7.72 m). Also note, however, that the range of observed values for all three *Event Type*s was wide (Figure 12). In addition, although the dissuasion-trigger response distances averaged close to the calibrated core-envelope trigger distance of 170 m, the averages for detections and warning signal triggers were notably shorter than the expected 240 m core-envelope trigger distances for those events.

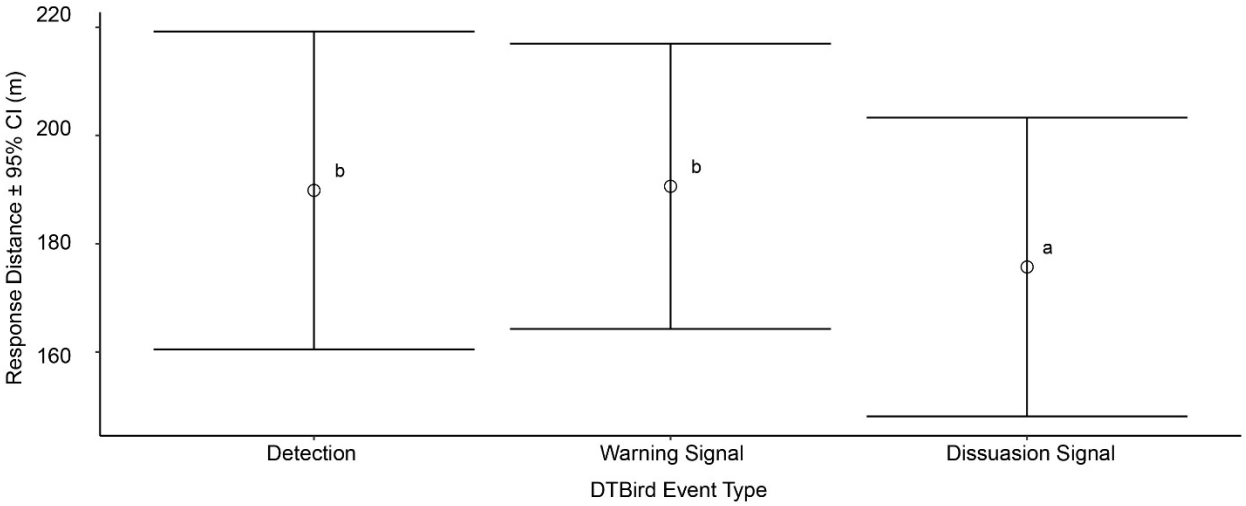

Figure 13. Modeled relationship between DTBird response distances and detection and deterrent-triggering event types, with shared letters indicating pairwise differences that are not significant at *P* ≤ 0.05.

***Sky Backdrop*:** Response distances and cloud cover were positively correlated, with the average response distance increasing with the progression from fair to overcast skies (Figure 13). Parameter tests and post-hoc comparisons of estimated marginal means confirmed that response distances averaged a significant 19.4 m shorter under fair skies (defined as few if any small clouds in the sky) than under overcast skies (defined as complete or near-complete, dense cloud cover with little to no penetration of blue sky or large sunspots), with the average responses under partly cloudy (defined as more than a few small clouds but <50% cloud cover) and mostly cloudy skies (≥50% up to near-complete cloud cover but with distinct patches of

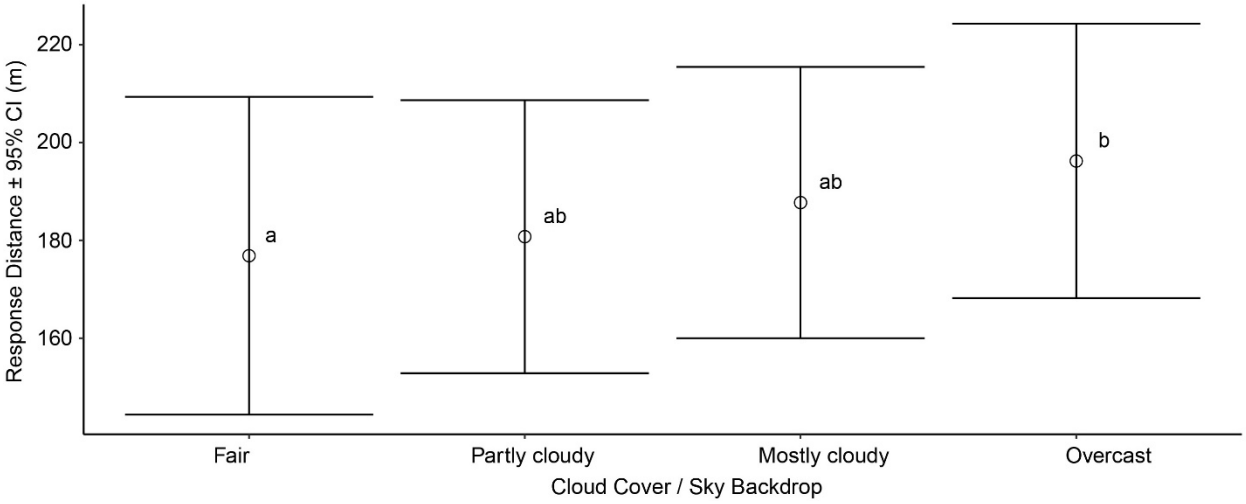

Figure 12. Modeled relationship between DTBird detection and deterrent-triggering response distances and sky backdrop / cloud cover categories, with shared letters indicating pairwise differences that are not significant at *P* ≤ 0.05.

blue and/or brighter clouds) intermediate in the progression and not significantly different from other categories.

***Ground Speed*:** Response distances tended to increase as the rate of UAV travel relative to fixed points on the ground increased (Figure 14).

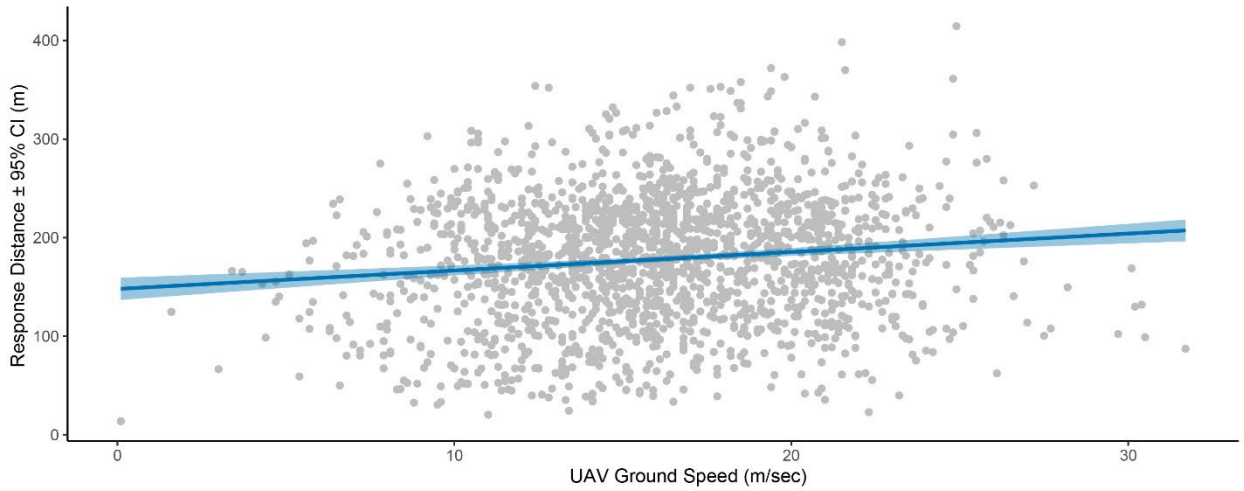

Figure 14. Modeled relationship (±95% confidence interval) between DTBird detection and deterrent-triggering response distances and UAV ground speed, or rate of travel relative to a fixed point on the ground, as measured by UAV avionics during sampling flights.

***Wind Speed*:** Response distances tended to increase as the wind speed—measured in flight by the UAV avionics—increased (Figure 15).

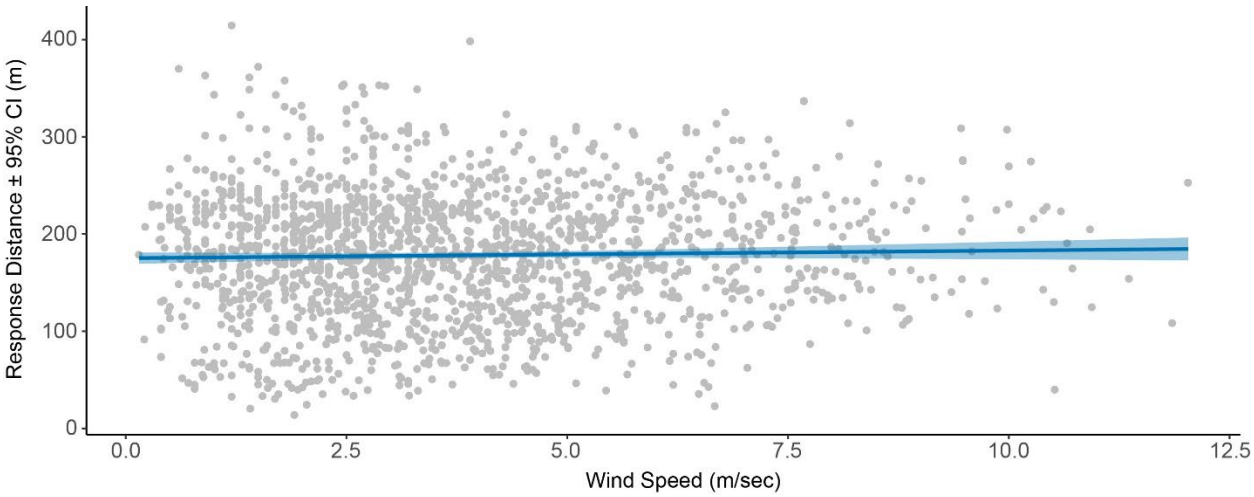

Figure 15. Modeled relationship between DTBird detection and deterrent-triggering response distances and wind speed as measured by UAV avionics during sampling flights.

***Roll Angle* : *Pitch Angle* Interaction:** The degree to which a UAV rolled to one side or the other or pitched up or down while in flight influenced DTBird response distances in an interactive manner (Figure 16). *Roll Angle* was shown to be the strongest predictor of the two variables (Table 4), with observed values ranging from approximately -59° (left roll) to +41° (right roll). The interactive influence of *Pitch Angle* (observed values from -20° pitched down to +36° pitched up) reflected that pitching and rolling often acted in concert to increase exposure of the UAV profile to the cameras, but concurrent maximization of both metrics was effectively impractical.

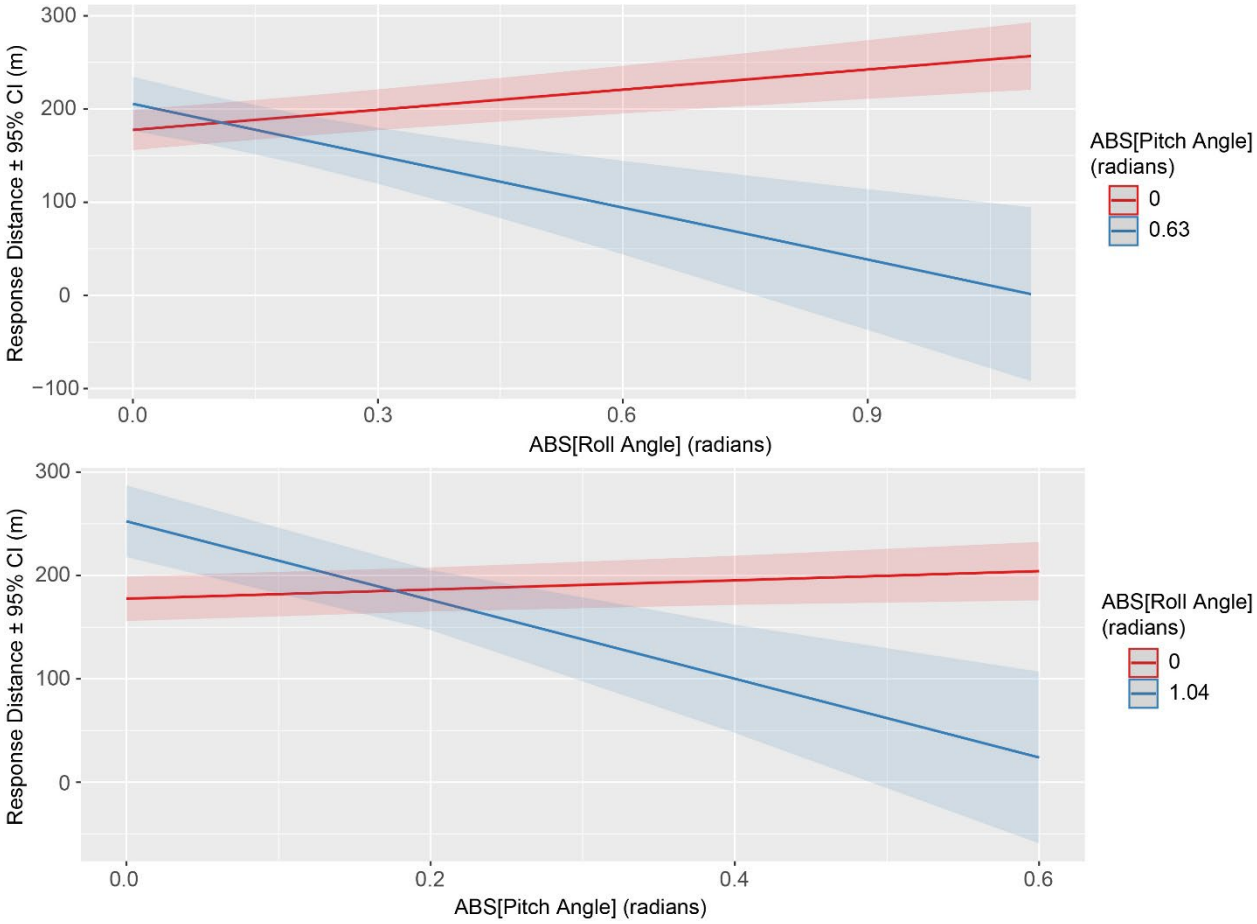

Figure 16. Modeled relationships between DTBird detection and deterrent-triggering response distances and the interactive influence of UAV pitch and roll angles.

More specifically, graphical illustrations of this interactive relationship indicated the following:

- With a low *Pitch Angle* (i.e., aircraft flying near nose-to-tail level), the more the UAV rolled from side to side (e.g., bouncing around in the wind or banking in a turn), the more the response distance increased.
- With a low *Roll Angle* (i.e., aircraft flying with wings near level), greater *Pitch Angle*s also tended to increase response distances to a lesser degree.

- Combinations of moderate pitch and roll angles were associated with moderate to moderately high response distances, but concurrent maximization of both stability metrics was effectively impractical, because it would translate to the aircraft stalling and falling out of the sky. Hence, the indications in Figure 16 that as one stability metric increased, the other generally declined, and vice versa, which was largely a result of the automated avionics programming explicitly striving to avoid stalling the aircraft.

4.5.2 *In Situ* Behavioral Responses of Eagles and Raptors (Objective 2)

Table 5 summarizes the classified large-raptor deterrence events from the two study sites that we analyzed for this assessment.

Table 55. DTBird events recorded from January through August 2017 at the Manzana Wind Power Project in California and from September 2021 through August 2022 at the Goodnoe Hills Wind Farm in Washington, which formed the basis for assessing the behavioral responses of eagles and other large raptors to DTBird audio deterrents.

	Manzana	Goodnoe Hills		
Species[1]	Deterrents Broadcasting	Deterrents Broadcasting	Deterrents Muted	Total
Golden Eagle	80	33	45	158
Bald Eagle	1	14	25	40
Unknown Eagle	0	11	9	20
Turkey Vulture	21	52	54	127
Buteo[2]	122	52	55	229
Golden Eagle or Vulture	39	7	3	49
Golden Eagle or Buteo	7	3	6	16
Unknown Eagle/Vulture	11	34	49	94
Unknown Eagle/Buteo	0	16	22	38
Total	281	222	268	771

[1] Classifications represent all cases where we either confirmed or strongly suspected ("probable") involvement of the relevant species or species group.

[2] Primarily red-tailed hawks year-round at both sites and rough-legged hawks during winter at the Goodnoe Hills.

4.5.2.1 Evaluating Differences in Behavioral Responses Between Sites

Given many cases where we could not confidently classify the species of raptor detected and tracked by the DTBird systems, we began our assessment by examining the deterrent response patterns reflected in all 503 of the selected cases involving large raptors exposed to broadcasted deterrents at the two study sites (Table 5). Overall, we classified 73% of the Manzana cases and 63% of the Goodnoe Hills cases as either confirmed or potentially effective responses (Table 6). The chi-square analysis of this dataset indicated a marginally significant difference ($0.05 < P \leq 0.10$) in the response patterns at the two sites ($\chi^2 = 5.59$, df = 2, $P = 0.061$). Post-hoc comparisons further indicated that the higher proportion of *Confirmed effective* responses approached significance only at the Manzana site ($P = 0.076$), the proportion of *Potentially effective* responses did not differ at the two sites ($P = 0.683$), and the proportion of *Ineffective* ($I = N + Z$) responses was marginally higher at the Goodnoe Hills ($P = 0.023$ falls

below the Bonferroni-corrected significance threshold for maintaining an overall Type II error rate of ≤0.10).

Table 66. Classification of the effectiveness of behavioral responses (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for all large raptors combined (eagles, vultures, and buteos) at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

Classified	Manzana		Goodnoe Hills	
Response	Number of Cases	%	Number of Cases	%
Confirmed Effective (*CE*)	118	42.0	76	34.2
Potentially Effective (*PE*)	87	31.0	69	29.3
Not Effective (*N*)	13	4.6	17	7.2
No Response (*Z*)	63	22.4	60	29.3
Total	281	–	222	–

Note: test of independence with N + Z lumped: χ2 = 5.59, df = 2, P = 0.061—indicating the overall pattern of responses was marginally different at the two sites. Bonferroni-corrected post-hoc comparisons confirmed a marginally higher proportion of Potentially effective responses and a marginally lower proportion of Ineffective (N+Z) responses at the Manzana site.

Focused on confirmed/probable golden eagles, the proportion of confirmed/potentially effective responses was again higher at the Manzana site (79%) compared to the Goodnoe Hills (60%) (Table 7), and the overall chi-square analysis again indicated that the pattern of variation among the *Response* classifications was at least marginally different at the two sites (χ^2 = 5.84, df = 2, *P* = 0.054). Post-hoc comparisons further indicated that the proportion of *Confirmed effective* responses was marginally higher at the Manzana site (*P* = 0.027), the proportion of *Potentially effective* responses did not differ at the two sites (*P* = 0.629), and the higher proportion of *Ineffective* responses at the Goodnoe Hills approached significance (*P* = 0.047).

Table 77. Classification of the effectiveness of behavioral responses (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable golden eagles at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

Classified	Manzana		Goodnoe Hills	
Response	Number of Cases	%	Number of Cases	%
Confirmed Effective (*CE*)	40	50.0	9	27.3
Potentially Effective (*PE*)	23	28.8	11	33.3
Not Effective (*N*)	3	3.7	5	15.2
No Response (*Z*)	14	17.5	8	24.2
Total	80	–	33	–

Note: chi-square test of independence with *N* + Z lumped: χ^2 = 5.84, df = 2, *P* = 0.054—indicating the overall pattern of responses was marginally different at the two sites. Bonferroni-corrected post-hoc comparisons confirmed a marginally higher proportion of *Confirmed effective* responses and a marginally lower proportion of *Ineffective* (N+Z) responses at the Manzana site.

For confirmed/probable turkey vultures, the proportion of confirmed/potentially effective responses was again higher at the Manzana site (81%) compared to the Goodnoe Hills site (61%) (Table 8), and the overall chi-square analysis indicated that the pattern of variation among the *Response* classifications differed at the two sites (χ^2 = 6.20, df = 2, *P* = 0.045). Post-hoc comparisons further indicated that the proportion of *Confirmed effective* responses was higher at the Manzana site (*P* = 0.015), the proportion of *Potentially effective* responses did not differ at the two sites (*P* = 0.424), and the higher proportion of *I* responses at the Goodnoe Hills approached significance (*P* = 0.069).

Table 88. Classification of the effectiveness of behavioral responses (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable turkey vultures at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

Classified Response	Manzana		Goodnoe Hills	
	Number of Cases	%	Number of Cases	%
Confirmed Effective (*CE*)	11	52.4	12	23.1
Potentially Effective (*PE*)	6	28.6	20	38.4
Not Effective (*N*)	0	0	4	7.7
No Response (*Z*)	4	19.0	16	30.8
Total	21	–	52	–

Note: Chi-square test of independence with N + Z lumped: χ2 = 6.20, df = 2, P = 0.045—indicating that the overall pattern of responses differed at the two sites. Bonferroni-corrected post-hoc comparisons confirmed a higher proportion of *Confirmed effective* responses at the Manzana site.

For confirmed/probable buteos, the difference between the overall proportions of confirmed/potentially effective responses was again notably higher at the Manzana site (72%) than at the Goodnoe Hills (56%). The chi-square analysis confirmed a significant difference in pattern at the two sites (χ^2 = 6.31, df = 2, *P* = 0.043; Table 9). Post-hoc comparisons further indicated that the proportion of *Confirmed effective* responses did not differ at the two sites (*P* = 0.095), but the proportion of *Potentially effective* responses was marginally higher (*P* = 0.028) and the proportion of *I* responses was marginally lower (*P* = 0.035) at the Manzana site.

Table 99. Classification of the effectiveness of behavioral responses (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable buteos at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

Classified Response	Manzana		Goodnoe Hills	
	Number of Cases	%	Number of Cases	%
Confirmed Effective (*CE*)	44	36.1	19	36.6
Potentially Effective (*PE*)	44	36.0	10	19.2
Not Effective (*N*)	8	6.6	5	9.6
No Response (*Z*)	26	21.3	18	34.6
Total	122	–	52	–

Note: Chi-square test of proportions: χ2 = 6.31, df = 2, P = 0.042—indicating the overall pattern of responses differed at the two sites. Bonferroni-corrected post-hoc comparisons confirmed a marginally higher proportion of Potentially effective responses and a marginally lower proportion of Ineffective (N + Z) responses at the Manzana site.

In relation to collision *Risk*, the raw percentage results for the multi-species Manzana dataset suggested that the proportion of *Confirmed effective* responses to broadcasted deterrents increased from 36% to 49% as the classified level of pre-exposure risk increased from low to high, whereas the proportions of *Potentially effective* and *I* responses each decreased by seven percentage points with increasing exposure risk (Table 10). In contrast, the multi-species Goodnoe Hills dataset suggested that the proportions of both *Confirmed effective* and *Potentially effective* responses were highest and the proportion of *I* responses lowest for birds at moderate pre-exposure risk.

Table 1010. Classification of the effectiveness of behavioral responses to broadcasted DTBird audio deterrents (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for all large raptors combined by site and classified risk level before deterrent exposure at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

	Site / Risk Level							
	Manzana				Goodnoe Hills			
Response	**Low**	**Med**	**High**	**Total**	**Low**	**Med**	**High**	**Total**
Confirmed Effective (*CE*)	42	58	18	118	28	39	9	76
Potentially Effective (*PE*)	40	37	10	87	27	31	7	65
Ineffective (*I = N + Z*))	36	31	9	76	40	29	12	81
Total Cases	118	126	37	281	95	99	28	222
% Confirmed Effective	36	46	49	42	29	39	32	34
% Potentially Effective	34	29	27	31	28	31	25	29
% Ineffective	31	25	24	27	42	29	43	36

The *Response–Risk* data for confirmed/probable golden eagles were sparse across many cells of the relevant 3 x 3 contingency tables for both sites, especially the Goodnoe Hills, which may limit the value of generated insight (Table 11). The Manzana data suggested that the proportions of *Confirmed effective* responses were higher for birds at high (50%) and especially moderate (58%) risk of exposure than for birds at low risk of exposure (40%), and the proportions of *I* responses were concomitantly lower for birds at moderate to high risk. In contrast, the Goodnoe Hills data showed a modest increasing trend in the proportions of *Confirmed effective* responses as risk increased (22–33%); however, among birds at moderate risk of exposure, the highest proportion (44%) exhibited relatively subtle *Potentially effective* responses, and the highest proportions of birds at both low (56%) and high (50%) risk of exposure exhibited no effective responses.

Table 1111. Classification of the effectiveness of behavioral responses to broadcasted DTBird audio deterrents (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable golden eagles by site and classified risk level before deterrent exposure at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

	Site / Risk Level							
	Manzana				Goodnoe Hills			
Response	**Low**	**Med**	**High**	**Total**	**Low**	**Med**	**High**	**Total**
Confirmed Effective (*CE*)	12	21	7	40	2	5	2	9
Potentially Effective (*PE*)	8	11	4	23	2	8	1	11
Ineffective (*I = N + Z*))	10	4	3	17	5	5	3	13
Total Cases	30	36	14	80	9	18	6	33
% Confirmed Effective	40	58	50	50	22	28	33	27
% Potentially Effective	27	31	29	29	22	44	17	33
% Ineffective	33	11	21	21	56	28	50	39

The Manzana sample sizes for confirmed/probable turkey vultures were sparse when broken out into a 3 x 3 *Response–Risk* table; however, the pattern of sparseness suggested that vultures at moderate to high risk of exposure exhibited a pronounced tendency to respond effectively, whereas birds at low risk of exposure were close to equally likely to exhibit any one

of the three responses (Table 12). In contrast, the Goodnoe Hills data suggested that *Confirmed effective* responses were least likely regardless of the pre-exposure risk level and were proportionately least common among birds at high risk, but no other consistent patterns were evident.

Table 1212. Classification of the effectiveness of behavioral responses to broadcasted DTBird audio deterrents (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable turkey vultures by site and classified risk level before deterrent exposure at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

	Site / Risk Level							
	Manzana				Goodnoe Hills			
Response	**Low**	**Med**	**High**	**Total**	**Low**	**Med**	**High**	**Total**
Confirmed Effective (*CE*)	4	6	1	11	5	5	2	12
Potentially Effective (*PE*)	5	0	1	6	9	6	5	20
Ineffective (*I = N + Z*))	4	0	0	4	7	9	4	20
Total Cases	13	6	2	21	21	20	11	52
% Confirmed Effective	31	100	50	52	24	25	18	23
% Potentially Effective	38	0	50	29	43	30	45	38
% Ineffective	31	0	0	19	33	45	36	38

For confirmed/probable buteos, neither of the site-specific datasets exhibited distinctive trends in the response patterns in relation to pre-exposure risk levels (Table 13). At the Manzana site, overall variation across cells of the 3 x 3 *Response–Risk* table was not pronounced. The highest proportion of birds at high risk (44%) exhibited *Confirmed effective* responses, whereas marginally highest proportions of the birds at low (40%) and moderate (36%) risk exhibited *Potentially effective* responses. At the Goodnoe Hills, the proportions of *I* responses were notably highest for birds at both low and high risk, whereas the proportion of *Confirmed effective* responses was notably highest for birds at moderate risk.

Table 1313. Classification of the effectiveness of behavioral responses to broadcasted DTBird audio deterrents (combined responses to warning and dissuasion signals acting alone or in tandem) in reducing collision risk for confirmed and probable buteos by site and classified risk level before deterrent exposure at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

	Site / Risk Level							
	Manzana				Goodnoe Hills			
Response	**Low**	**Med**	**High**	**Total**	**Low**	**Med**	**High**	**Total**
Confirmed Effective (*CE*)	15	21	8	44	8	9	2	19
Potentially Effective (*PE*)	17	22	5	44	2	7	1	10
Ineffective (*I = N + Z*))	11	18	5	34	16	4	3	23
Total Cases	43	61	18	122	26	20	6	52
% Confirmed Effective	35	34	44	36	31	45	33	37
% Potentially Effective	40	36	28	36	8	35	17	19
% Ineffective	26	30	28	28	62	20	50	44

The performance standard of ≥50% successful or effective deterrence for golden eagles established based on the initial Manzana pilot study (H. T. Harvey & Associates 2018) was

further corroborated for that site by the initial 53% estimate derived from the subsequent expansion of that site-specific assessment to include a full year of data. Further minor adjustments to the relevant dataset in preparation for the multi-site evaluation presented herein modified that estimate to 50% *Confirmed effective* responses, with another 29% *Potentially effective* responses, yielding a total estimated probable effectiveness of 79% for golden eagles (Table 14). In comparison, the Goodnoe Hills results indicated a lower 27% confirmed effective responses, falling well below the established performance standard; however, the combined estimate of 60% confirmed/probable effective responses, though still notably lower than at the Manzana site, did exceed the 50% performance threshold. Similar patterns were shown for vultures and the multi-species group, except that the proportion of effective responses for the multi-species group fell below the 50% threshold. In contrast, for buteos the proportions of effective responses did not differ at the two sites and were well below the 50% threshold (27–29%); however, the combined proportion of confirmed/probable effective responses was again notably higher at the Manzana site (72%) than at the Goodnoe Hills site (56%) (Table 14).

Table 1414. Percentages of behavioral responses to broadcasted DTBird deterrents (combined responses to warning and dissuasion signals acting alone or in tandem) classified as effective or potentially effective in reducing collision risk for different species groups at the Manzana Wind Power Project in California and the Goodnoe Hills Wind Farm in Washington.

Species Group	Manzana	Goodnoe Hills
Golden Eagles	50 / 79[1]	27 / 60
Vultures	52 / 81	23 / 61
Buteos	36 / 72	37 / 56
All Groups Combined	42 / 73	34 / 63

[1] First number = % of responses confirmed effective; second number = overall % of confirmed + potentially effective responses.

4.5.2.2 Evaluating Factors Influencing Behavioral Responses to Deterrents

Given that the initial chi-square analyses pointed to at least marginally significant differences in the deterrence response patterns of golden eagles and other large raptors at the two study sites, we did not consider pursuing Objective C as outlined in the Introduction. Instead, we pursued the second element of Objective B, which entailed preparing LGLM analyses to provide further insight about potential drivers of the evident site-specific differences in the apparent sensitivity of raptors to the broadcasted deterrents.

Multi-species Model: The LGLM analysis based on the multi-species dataset resulted in the final model listed below (and see Table 15) and the interpretations that follow:

> Log(Odds of effective deterrence) ~ *Site* + *Species Group* + *Preexposure Risk* + *Wind Speed* + *Species Group* * *Wind Speed*

Diagnostics for this final model revealed no outliers and residuals consistent with adequate model fit.

Table 1515. Comparison of AIC scoring results for top candidates and selected other multi-species logistic GLMs portraying potential relationships between the probability of effective deterrence and various predictors.

Candidate Model[1]	AIC[2]	ΔAIC	*McFadden's R*[2]
Site + Species Group + Preexposure Risk + Wind Speed + Species Group : Wind Speed	465.52	0.00	0.055
Site + Species Group + Preexposure Risk + Wind Speed + Species Group : Wind Speed + Site : Wind Speed	466.44	0.92	0.057
Site + Species Group + Wind Speed + Species Group : Wind Speed	466.87	1.35	0.044
Site	469.29	3.77	0.018
Site + Species Group	470.37	4.85	0.024
Site + Species Group + Preexposure Risk + Wind Speed + Species Group : Preexposure Risk + Species Group : Wind Speed + Site : Wind Speed	470.53	5.01	0.066
Site + Wind Speed	471.16	5.64	0.018
Site + Species Group + Wind Speed	471.92	6.40	0.025
Species Group*Wind Speed	474.30	8.78	0.024
Null model	475.60	10.08	–
Site + Species Group + Preexposure Risk + Wind Speed +Species Group : Site + Species Group : Preexposure Risk + Species Group : Wind Speed +Site : Preexposure Risk + Site : Wind Speed	477.23	11.71	0.068
Site + Species Group + Preexposure Risk + Wind Speed +Species Group : Site + Species Group : Preexposure Risk + Species Group : Wind Speed + Site : Wind Speed	477.33	11.81	0.068
Site + Species Group + Preexposure Risk + Wind Speed +Species Group : Site + Species Group : Preexposure Risk + Species Group : Wind Speed + Site : Preexposure Risk + Site : Wind Speed + Preexposure Risk : Wind Speed	481.22	15.70	0.068

[1] *Site* = Manzana or Goodnoe Hills wind facility. *Species Group* = eagle, vulture or buteo. *Preexposure Risk* (of approaching rotor swept area of spinning turbine prior to deterrent triggering) = low, moderate or high. *Wind Speed* measured at turbine in meters / second.

[2] Akaike Information Criterion score.

Table 1616. Parameters of final multi-species logistic GLM selected to represent relationship between the ln(odds of effective deterrence) and various predictors at the Manzana and Goodnoe Hills wind-energy facilities.

Parameter[1]	Estimate	SE	*z*	*P*
Intercept	0.6394	0.5112	1.251	0.211
Site–Manzana	0.7416	0.2439	3.041	0.002
Species Group–Eagle	-0.8740	0.5548	-1.575	0.115
Species Group–Vulture	-0.6512	0.6965	-0.935	0.350
Preexposure Risk–Low	-0.2023	0.3355	-0.603	0.547
Preexposure Risk–Moderate	0.3748	0.3395	1.104	0.270
Wind Speed	-0.0725	0.0508	-1.427	0.153
Species Group–Eagle : Wind Speed	0.2220	0.0858	2.587	0.010
Species Group–Vulture : Wind Speed	0.1562	0.0993	1.574	0.116

[1] *Site* reference category = Goodnoe Hills. *Species Group* reference category = buteo. *Preexposure Risk* reference category = high.

Site effect (*P* =0.002; Table 16) reflected a higher average probability of effective deterrence at the Manzana site (Figure 17).

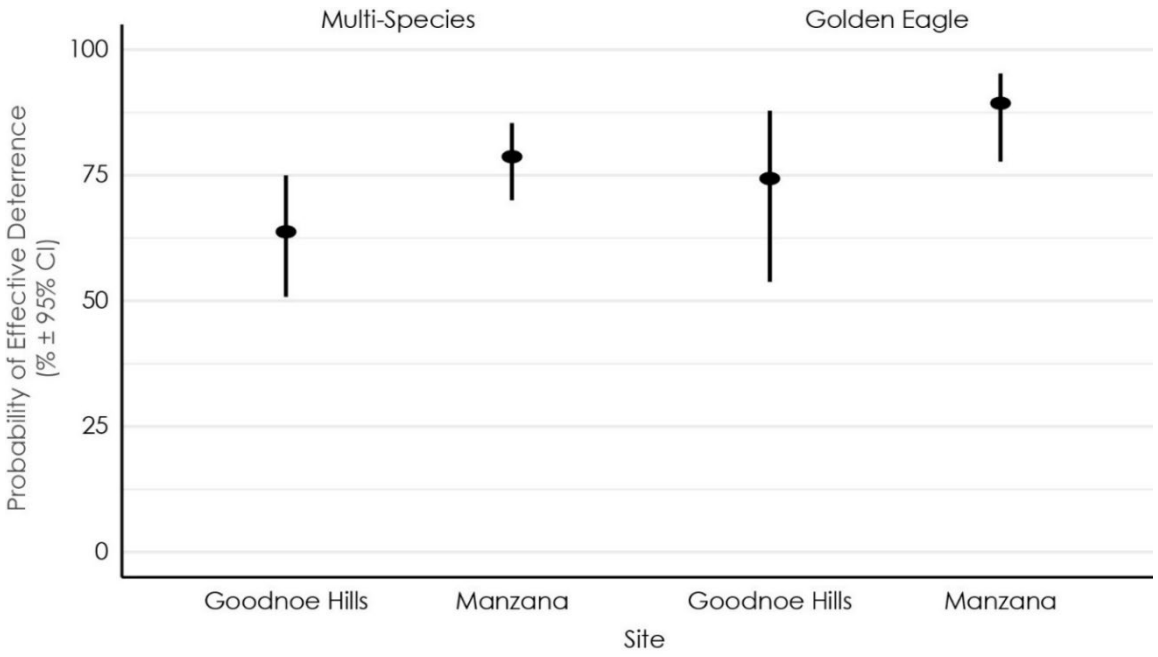

Figure 17. Modeled probability of effective DTBird deterrence for all large raptors combined and golden eagles alone at the two wind facilities evaluated in this study.

Preexposure Risk was only marginally significant (*P* = 0.069), but its inclusion reduced the AIC score by 1.35 points (Table 15). Birds facing moderate risk were the most likely to show effective deterrence responses, while birds facing low risk were the least likely to show effective responses; however, none of the pairwise differences were significant on their own, suggesting a gradient of variation rather than a discrete segregation of probability groups (Table 16, Table 20).

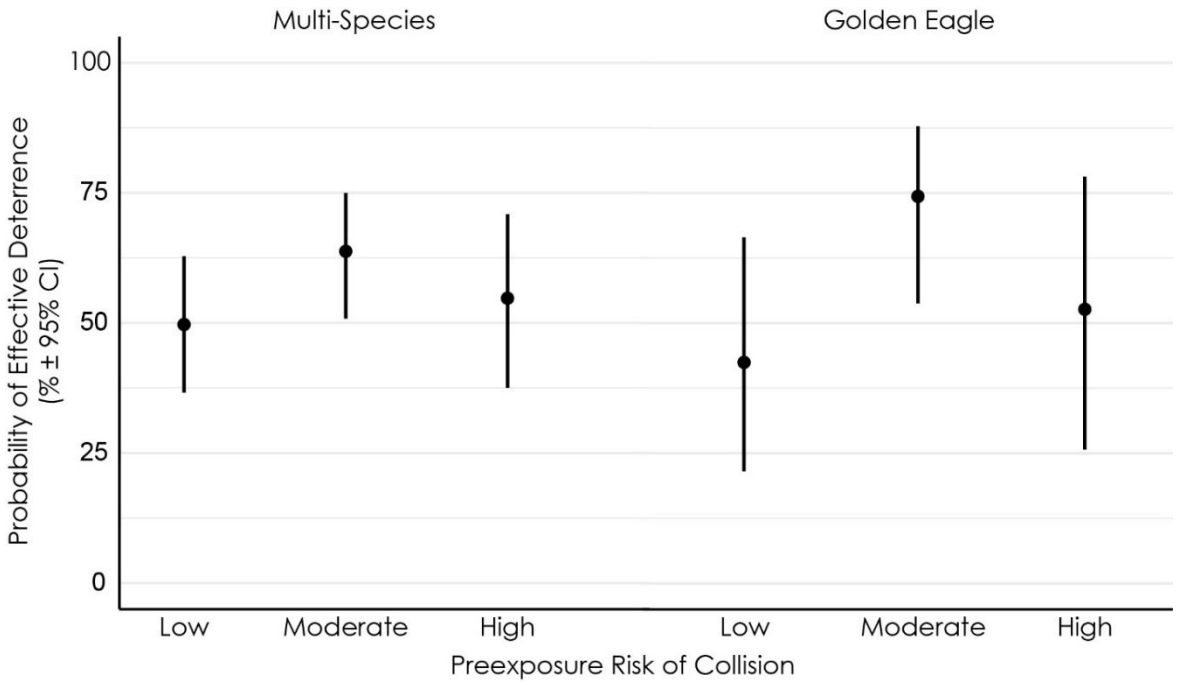

Figure 19. Modeled probability of effective DTBird deterrence for all large raptors combined and golden eagles alone in relation to classified risk of exposure to turbine collisions at the two wind facilities evaluated in this study.

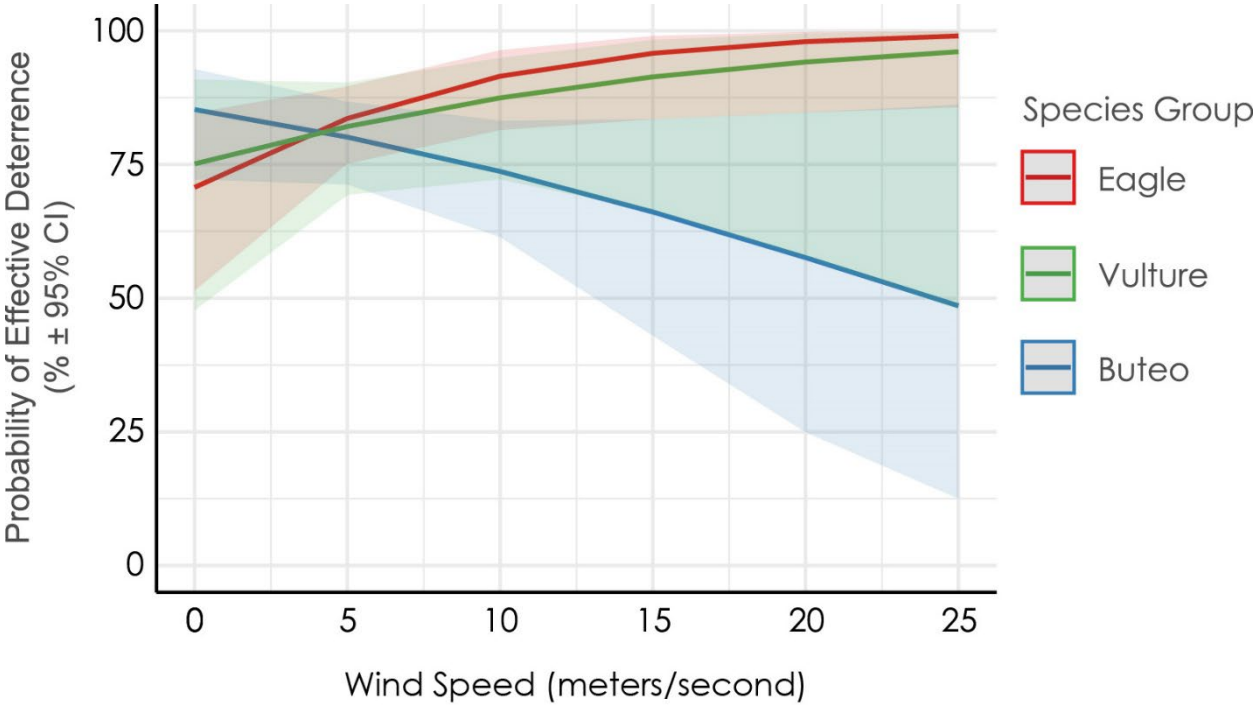

Figure 18. Modeled probability of effective DTBird deterrence for large raptors by species group and in relation to wind speed measured by turbine anemometer at time of events at the two wind facilities evaluated in this study.

Species Group and *Wind Speed* did not contribute significant main effects, but their 2-way

interaction was significant (*P* = 0.019). The *Species Group* * *Wind Speed* interaction reflected the following (Table 16, Figure 19):

- At low wind speeds below approximately 4 meters/second (m/s) (just above the turbine cut-in speed of 3 m/s), the probability of effective deterrence was lowest for eagles, slightly higher for vultures, and slightly higher still for buteos, whereas wind speeds above 4 m/s resulted in the opposite pattern.
- At wind speeds above approximately 4 m/s, the probability of effective deterrence was:
    - highest for eagles and increased strongly as wind speeds increased.
    - second highest for vultures and increased moderately as wind speeds increased.
    - lowest for the smaller buteos and decreased moderately as wind speeds increased.

Golden Eagle Model: The LGLM analysis for golden eagles resulted in the final model listed below (and see Table 17) and the interpretations that follow:

Log(Odds of effective deterrence) ~ *Site* + *Preexposure Risk* + *Wind Speed*

Diagnostics for this final model revealed no influential outliers and residuals consistent with adequate model fit.

Site effect (*P* =0.029; Table 18) reflected a higher average probability of effective deterrence at the Manzana site (Figure 17).

Preexposure Risk effect (*P* = 0.041) reflected that the probability of effective deterrence was highest for birds at moderate risk, moderate for birds at high risk, and significantly lowest for birds at low risk (Table 18, Figure 18).

Wind Speed was only marginally significant (*P* = 0.087; Table 18), but its inclusion reduced the AIC score by 1.2 pts (Table 17) and reflected a positive relationship with the probability of deterrence (Table 18, Figure 19).

Another model including the *Site* * *Wind Speed* interaction scored lowest on the AIC scale, but improved the AIC score by only a nominal 0.45 points compared to the second-best model chosen as the final. Further, the parameter-test *P* value for the interaction (0.118) exceeded even the *P* ≤0.10 threshold for marginal significance. Nevertheless, the suggested interactive relationship indicated a potentially interesting pattern, whereby (a) the probability of deterrence rose more quickly as wind speed increased at the Goodnoe Hills than at the Manzana site, and (b) as a consequence, was higher at the Manzana site at winds speeds below about 7 m/s, but was higher at the Goodnoe Hills at wind speeds greater than that (Figure 20).

Table 1717. Comparison of AIC scoring results for top candidates and selected other logistic GLMs portraying potential relationships for golden eagles between the ln(odds of effective deterrence) and various predictors.

Candidate Model[1]	AIC[2]	ΔAIC	*McFadden's R*[2]
Site + Preexposure Risk + Wind Speed + Site : Wind Speed	126.23	0.00	0.127
Site + Preexposure Risk + Wind Speed	126.68	0.45	0.108

Site + Preexposure Risk	127.90	1.67	0.083
Site + Wind Speed + Site : Wind Speed	128.64	2.41	0.078
Preexposure Risk + Wind Speed	129.47	3.24	0.071
Site + Preexposure Risk + Wind Speed + Site : Preexposure Risk + Site : Wind Speed	130.12	3.89	0.127
Site	131.01	4.78	0.029
Preexposure Risk	131.06	4.83	0.044
Wind Speed	131.92	5.69	0.022
Preexposure Risk + Wind Speed + Preexposure Risk : Wind Speed	132.51	6.28	0.079
Null model	132.51	6.28	–
Site + Preexposure Risk + Wind Speed + Site : Preexposure Risk + Site : Wind Speed + Preexposure Risk : Wind Speed	132.79	6.56	0.128

[1] *Site* = Manzana or Goodnoe Hills wind facility. *Species Group* = eagle, vulture or buteo. *Preexposure Risk* (of approaching rotor swept area of spinning turbine prior to deterrent triggering) = low, moderate or high. *Wind Speed* measured at turbine in meters / second.

[2] Akaike Information Criterion score.

Table 1818. Parameters of final logistic GLM selected to represent relationship between the ln (odds of effective deterrence) for golden eagles and various predictors at the Manzana and Goodnoe Hills wind-energy facilities.

Parameter[1]	Estimate	SE	*z*	*P*
Intercept	-0.6933	0.7694	-0.901	0.3675
Site–Manzana	1.0615	0.4867	2.181	0.0292
Preexposure Risk–Low	-0.4103	0.6253	-0.656	0.5118
Preexposure Risk–Moderate	0.9581	0.6470	1.481	0.1386
Wind Speed	0.1612	0.0942	1.711	0.0870

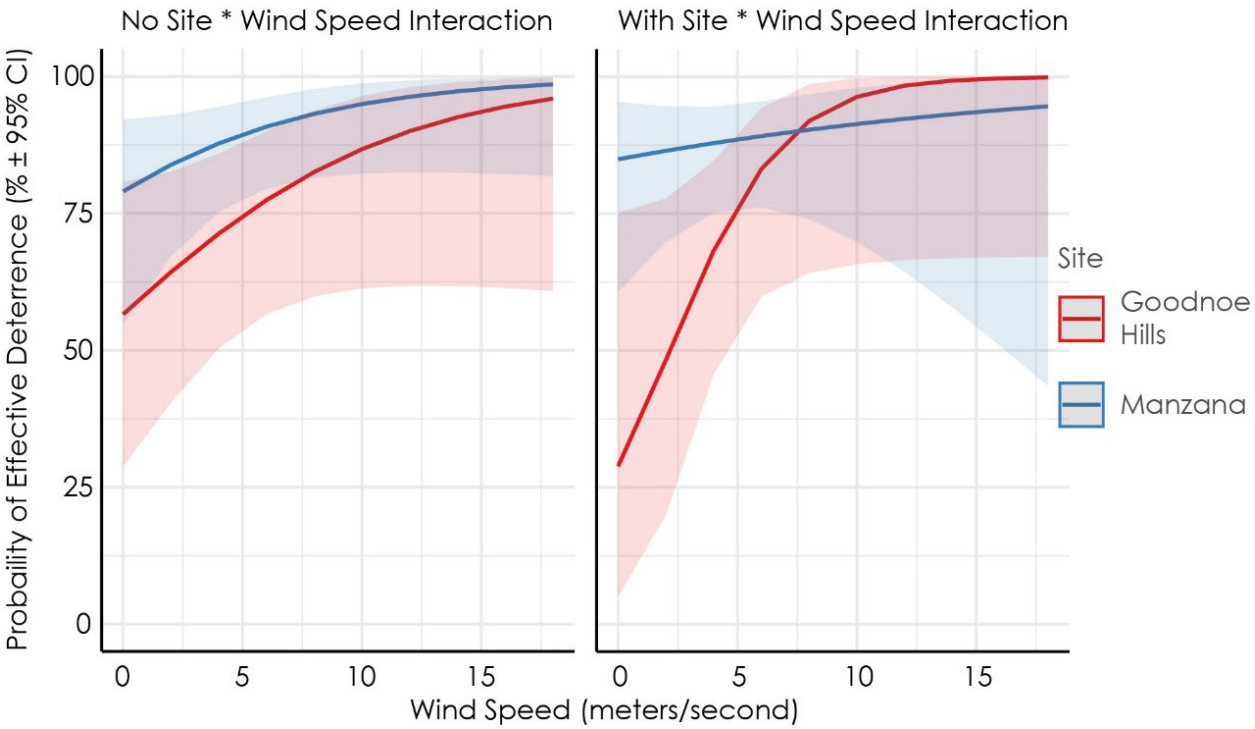

Figure 20. Modeled probability of effective DTBird deterrence for golden eagles in relation to wind speed measured by turbine anemometer at time of events at the two wind facilities evaluated in this study, showing results with and without *Site * Wind Speed* interaction (improves AIC score but nonsignificant *P* = 0.118 parameter test).

The final model and the model including the *Site * Wind Speed* interaction had a McFadden's pseudo-R^2 values of 0.108 and 0.127, respectively, and were the two models with the highest such values (Table 17). The closeness of the pseudo-R^2 values of these two models indicates that they have essentially equal ability to explain variation in deterrence probabilities. Both values are between 0.1 and 0.2, indicating "good" predictive value (values of 0.1–0.2 are considered a "good" result, while values of 0.2–0.4 are considered an "excellent" result; McFadden 1974, 1979).

4.5.3 Factors Influencing False Positive Detection Rates (Objective 3)

4.5.3.1 DTBird Event Classifications

The 10-month, seven-turbine dataset analyzed from the Manzana site to derive results for this multi-site assessment involved 3,051 detections that triggered one or both deterrents (i.e., warning and/or dissuasion signals). With unknown big birds, unknown medium/large raptors, and unknown birds proportionately allocated where appropriate to the large raptors and NTAFP groups, the Manzana records included 789 detections classified as large soaring raptors, 917 detections classified as TFPs, and 1,212 detections classified as NTAFPs (Table 19). The analyzed 11-turbine dataset from Year 1 at the Goodnoe Hills involved 11,265 detections that triggered deterrents, including 1,529 classified as relevant raptors, 5,744 as TFPs, and 3,955 as NTAFPs. The analyzed intermittently 14-turbine dataset from Year 2 at the Goodnoe Hills

involved 8,075 detections that triggered deterrents, including 1,673 classified as relevant raptors, 3,441 as TFPs, and 2,958 as NTAFPs.

At Manzana, NTAFPs caused an estimated 40% of all deterrent triggers, TFPs caused 30%, large raptors caused 26%, and birds that remained classified as unknown medium/large raptors caused 4%. Particularly high raven activity at one DTBird turbine contributed to complaints from a residence approximately 500 meters away from that turbine. At Goodnoe Hills, adjusted NTAFPs caused a similar 36% of all deterrent triggers, whereas TFPs caused a higher 48% and large raptors caused a lower 17% of the total. Confirmed common ravens caused 24% of all false-positive deterrent triggers at Goodnoe Hills and 15% at Manzana.

Table 1919. DTBird Detection Events that Triggered Deterrents Classified as Large Raptors, True False Positives (TFPs), and Nontarget Avian False Positives (NTAFPs) at the Manzana Wind Power Project in California and Goodnoe Hills Wind Farm in Washington

				Large Raptors[2]		TFPs[3]		NTAFPs[4]	
Site	Number of Operational DTBird Systesms	Period of Record	Total Detection Events[1]	Number of Detection Events	Average Events/ Turbine/ Day	Number of Detection Events	Average Events/ Turbine/ Day	Number of Detection Events	Average Events/ Turbine/ Day
Manzana	7	Jan–Oct 2017	3,051	789	1.1	917	1.3	1,212	1.7
Goodnoe Hills Year 1	11	Sep 2021–Aug 2022	11,260	1,529	1.3	5,744	4.9	3,955	3.3
Goodnoe Hills Year 2	14	Sep 2022–Jul 2023	8,075	1,673	1.5	3,441	3.0	2,958	2.6
Total	Max 21	–	22,386	3,991	1.3	10,102	3.3	8,125	2.7

[3] Includes unidentified medium/large raptors that we did not reclassify as Large Raptors or NTAFPs and were excluded from analyses.
[4] Restricted to large soaring species; i.e., eagles, vultures, buteos, harriers, and ospreys.
[5] Includes events triggered by inanimate objects, insects, and software/video interpretation errors and failures.
[6] Includes events triggered by birds other than large soaring raptors and unknown medium/large raptors.

6.5.3.2 True False Positives

At Goodnoe Hills, the additional false-positive filtering adjustments made in January 2023 reduced the overall rate of TFP deterrent triggers from approximately 529 to 71 per month across all sampled turbines (87% reduction). Substantial proportional reductions in the monthly TFP deterrent triggering rates included those caused by insects (97%), sky artifacts (94%), floating debris (93%), other turbine equipment features (91%), spinning turbine blades (88%), precipitation (67%), and software/video issues (39%). Note, however, that unequal seasonal sampling and variation also could have affected the outcomes for insects, sky artifacts, floating debris, and precipitation. In addition, modifications of the absolute numbers substantially altered the proportional contributions of different types of TFPs observed at Goodnoe Hills in only a few cases. The proportion of blade-related TFPs declined only slightly from 32% of all TFP deterrent triggers in Year 1 to 28% post-adjustments in Year 2. The proportion of insect-related TFPs declined more substantially from 28% in Year 1 to 9% post-adjustments in Year 2, and the proportion of sky artifact TFPs declined from 23% in Year 1 to 9% post-adjustments in

Year 2. Concomitantly, the proportion of TFPs caused by aircraft increased from 11% in Year 1 to 30% post-adjustments in Year 2, and the proportion of TFPs caused by software failures increased from 4% in Year 1 to 18% post-adjustments in Year 2.

The range of TFP source types was similar but the percentage contributions of different sources varied at the two study sites (Attachment 6: Table 2). Before the false-positive filtering was adjusted at the Goodnoe Hills study, turbine blades (30–32% of TFPs) and insects (28–48%) variably ranked as the most and second-most common sources of TFPs, with TFPs caused by aircraft (6–11%) and sky artifacts (9–23%) variably ranked as the third and fourth most common sources. At Manzana by contrast, aircraft caused a majority of the TFPs (60%), sky artifacts caused the second highest proportion (25%), and insects caused a notably lower, third highest proportion (5%). The only other instance where another source caused more than 5% of the TFPs recorded during one of the four site-sampling periods involved software failures during the Goodnoe Hills Year 2 post-adjustments period (18% of TFPs in that period).

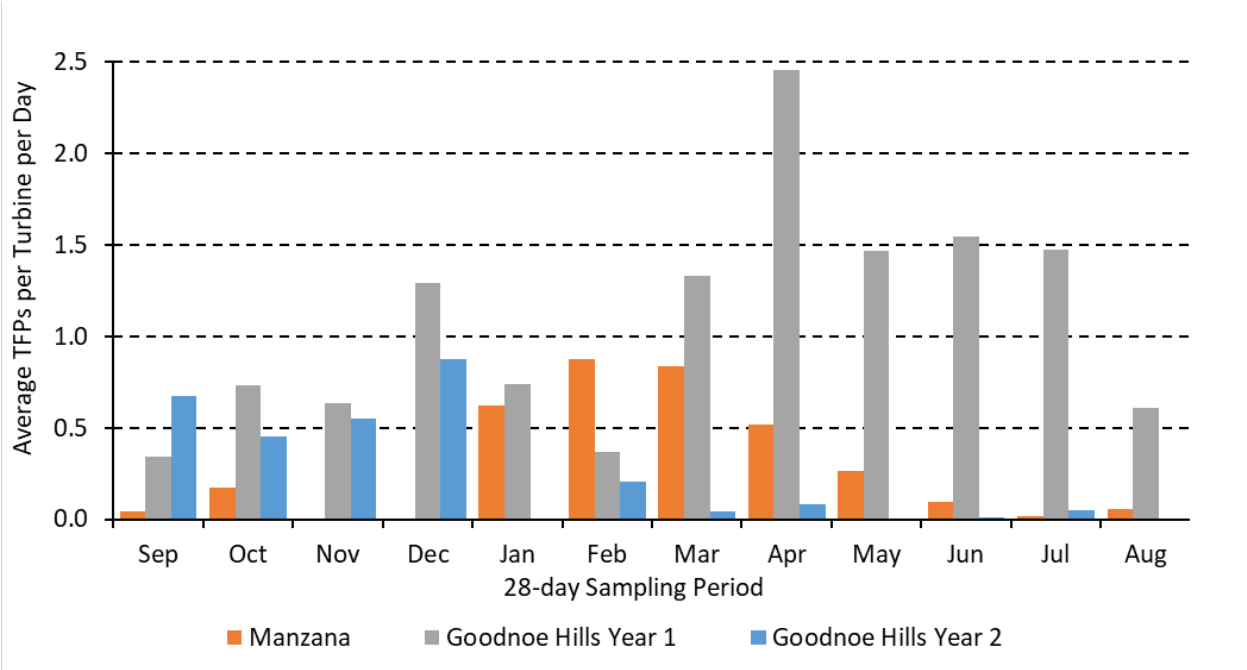

Figure 21. Rates of True False Positives Caused by Insects that Triggered DTBird Deterrents by Month at the Manzana Wind Power Project in California (January – October 2017) at the Goodnoe Hills Wind Farm in Washington (September–August 2021–2022 and 2022–2023).

The proportion of TFPs caused by insects showed distinctly different patterns both between years at Goodnoe Hills and between the two sites (Figure 21). At Manzana, insect TFPs were generally much less prevalent than at Goodnoe Hills and occurred mostly in early to mid-summer. During Goodnoe Hills Year 1, insect TFPs started out high in the fall, were largely absent during winter, began to ramp up in spring, and peaked in summer. In contrast, during Goodnoe Hills Year 2, insect TFPs were very high initially during fall (expanding the summer peak from Year 1), dropped off and again were rare through winter, but unlike during Year 1, remained low and comparable to the Manzana rates after that.

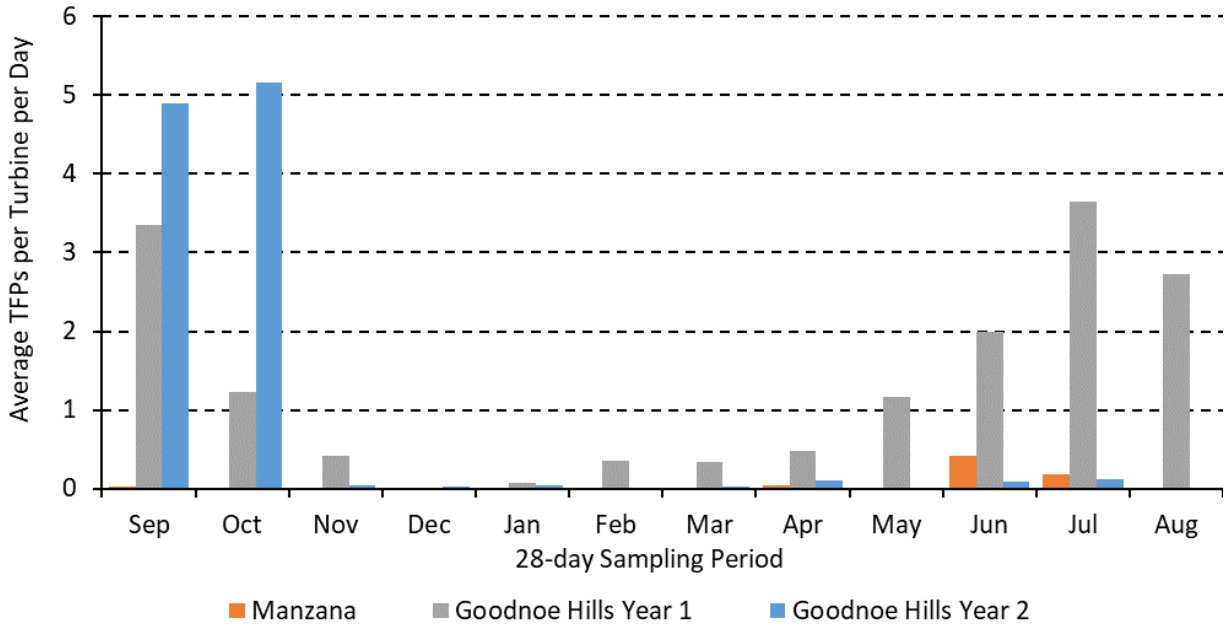

Figure 22. Rates of True False Positives Caused by Sky Artifacts that Triggered DTBird Deterrents by Month at the Manzana Wind Power Project in California (January–October 2017) and at the Goodnoe Hills Wind Farm in Washington (September–August 2021–2022 and 2022–2023).

The prevalence of TFPs caused by sky artifacts showed very different patterns across *28d Cycles* in Years 1 and 2 at Goodnoe Hills, whereas the patterns were much more similar for Manzana and Goodnoe Hills Year 2 (Figure 22). After the fifth cycles, sky artifact TFPs dropped off markedly and remained low at both the Manzana site and at Goodnoe Hills during Year 2. Note that, while this drop-off marked the time when further changes were made in the false positive filtering algorithms at Goodnoe Hills, it did not correspond to any such change at Manzana. After this point, though showing comparable rates and variation through the first 4–5 cycles, the rate of sky artifact TFPs increased markedly during Goodnoe Hills Year 1 and remained high through the 12th cycle, before dropping back down again to a moderate level during the 13th cycle (Figure 22, noting that for Goodnoe Hills the indicated patterns across months are essentially the same as for *28d Cycles*, whereas *28d Cycle* 1 was in January at Manzana). Considering the patterns in relation to calendar months further suggested that seasonal variation in the relative prevalence of sky artifact TFPs also might have contributed to the observed patterns. Though temporal mismatches in the site-specific datasets confound seasonal comparisons, it appeared that sky artifact TFPs were most common at Manzana in late winter early spring and dropped off during summer, whereas the Goodnoe Hills Year 1 data suggested comparatively high rates across the year and an extended period of peak activity from spring through summer (Figure 22).

6.5.3.3 Nontarget Avian False Positives

The range of general categories of NTAFP sources was similar at the two study sites. The only material difference in the proportional representations was that the percentage of confirmed

common ravens was lower at Manzana (28% of classified NTAFPs) than during either sampling year at the Goodnoe Hills (39–42%), whereas the proportion of unidentified big birds that we ultimately classified as NTAFPs was higher at Manzana (57%) than it was during both years at Goodnoe Hills (40–45%) (Attachment 6: Table 3).

4.5.3.4 False Positive Deterrent Triggering Rates and Durations

The overall average large-raptor deterrent triggering rates were relatively consistent across the three primary site-sampling periods, ranging from 1.3–1.5 detections with deterrent triggers/turbine/day (Table 19). The overall average TFP deterrent triggering rates were more variable, ranging from a low of 1.3 detections with deterrent triggers/turbine/day at Manzana to a high of 4.9 detections with deterrent triggers/turbine/day during Year 1 at the Goodnoe Hills; the Year 2 TFP deterrent triggering rate at the Goodnoe Hills was midday between the other two estimates. The same general pattern of differences was evident among the NTAFP deterrent triggering rates (Table 20).

Table 2020. Overall Durations and Average Per Turbine Duration Rates for DTBird Deterrent Signals Triggered by True False Positives (TFPs) and Nontarget Avian False Positives (NTAFP) at the Manzana Wind Power Project in California and Goodnoe Hills Wind Farm in Washington.

		Warning Signals			Dissuasion Signals		
Site	Sampling Period	Number of Triggers	Total Duration (minutes)	Average Duration/ Turbine/Day (seconds)	Number of Triggers	Total Duration (minutes)	Average Duration/ Turbine/Day (seconds)
TFPs							
Manzana	10 months	487	294	17.3	662	370	33.5
Goodnoe Hills	Year 1	654	217	27.0	4820	2465	30.7
	Year 2 – 4.5 months pre-adjustments	493	78	26.5	2551	1361	32.0
	Year 2 – 6.5 months post-adjustments	199	589	23.4	685	383	33.6
NTAFPs							
Manzana	10 months	979	364	22.3	458	223	29.1
Goodnoe Hills	Year 1	2510	1097	26.2	173.5	960	33.2
	Year 2 – pre	1138	484	25.5	797	438	33.0
	Year 2 – post	1083	458	25.4	602	321	32.0

Standardized for variable sampling intensity, the overall average TFP-caused warning signal durations on turbine-days when deterrents were triggered averaged 17.3 seconds/turbine/day at the Manzana site and a significantly higher 26.2 seconds/turbine/day at the Goodnoe Hills site (Table 20). The average warning signal duration rate at the Goodnoe Hills declined from 27.0 seconds/turbine/day during Year 1 down to 23.4 seconds/turbine/day during the Year 2

post-adjustments period, but still remained notably longer than at Manzana. In contrast, the average duration rates for dissuasion signals rose slightly at Goodnoe Hills between Year 1 (30.7 seconds/turbine/day) and the Year 2 post-adjustments period (33.6 seconds/turbine/day), but in this case the higher Year 2 post-adjustments rate more closely matched the Manzana rate (33.5 seconds/turbine/day).

Similar patterns of variation were evident in the overall average NTAFP-caused deterrent signal duration rates (Table 20), except differences among the Goodnoe Hills sampling periods and between the two study sites were less pronounced, and the duration rates declined slightly for both warning and dissuasion signals between Year 1 and the post-adjustments Year 2 period at the Goodnoe Hills.

With the analysis limited to comparing results across 12 common 28d Cycles, the numbers of days from which samples were drawn to compose the GLMM relating daily turbine-specific counts of TFPs that triggered deterrents to Year and 28d Cycle at the Goodnoe Hills varied from 10–119 per turbine across 11 sampled turbines in Year 1, and from 57–97 per turbine across 14 sampled turbines in Year 2 (Table 21). For the analysis comparing Goodnoe Hills results by Year and Month, we excluded May from the comparison due to an absence of data from that month in Year 2. For this reason, the sample sizes used to compare Year 1 and Year 2 by Month at the Goodnoe Hills were slightly lower for Year 1 than in the 28d-Cycle analysis (Table 21). The GLMM relating daily turbine-specific TFP counts to Year and 28d Cycle revealed a highly significant main effect for Year (Wald $\chi 2$, $P < 0.0001$), a non-significant main effect for 28d Cycle ($P = 0.98$), and a highly significant interaction term ($P < 0.0001$). Nakagawa's marginal pseudo-R2 for the model was 0.288, indicating that the fixed effects in the model provided moderate explanatory power (Nakagawa and Schielzeth 2013). Given the significant interaction, we conducted planned post-hoc comparisons to identify significant pairwise differences between Years within 28d Cycles and among 28d Cycles within Years. These comparisons confirmed the substantial shift in TFP prevalence after the additional filtering adjustments were made during the fifth 28d Cycle of Year 2 (Figure 23). Before that, the TFP rates did not differ markedly during corresponding 28d Cycles of the two sampling years. After that, the TFP rates remained significantly lower in Year 2 than in Year 1 during all subsequent 28d Cycles. Further, the post-adjustments Year 2 rates remained consistently low post- adjustments, whereas the corresponding Year 1 rates rose steadily after the sixth cycle to the highest rate for the year during the twelfth cycle.

Table 2121. Numbers of turbine-specific days from which samples were drawn for investigating temporal differences in DTBird false-positive detection rates between sampling years at the Goodnoe Hills Wind Farm in Washington.

	Analysis by 28d Cycles: Common Cycles 1–12			Analysis by Month: All Months Except May		
Turbine	**Year 1**	**Year 2**	**Total**	**Year 1**	**Year 2**	**Total**
G29	–	95	95	–	95	95
G34	98	91	189	104	91	195
G35	89	87	176	88	87	175
G44	107	79	186	103	79	182

G45	108	87	195	103	87	190
G48	112	57	169	110	57	167
G49	105	90	195	101	90	191
G51	–	80	80	–	80	80
G56	–	70	70	–	70	70
G58	117	97	214	115	97	212
G59	104	57	161	106	57	163
G64	119	91	210	118	91	209
G67	112	86	198	111	86	197
G75	10	75	85	10	75	85
Total	1,081	1,142	2,223	1,069	1,142	2,211

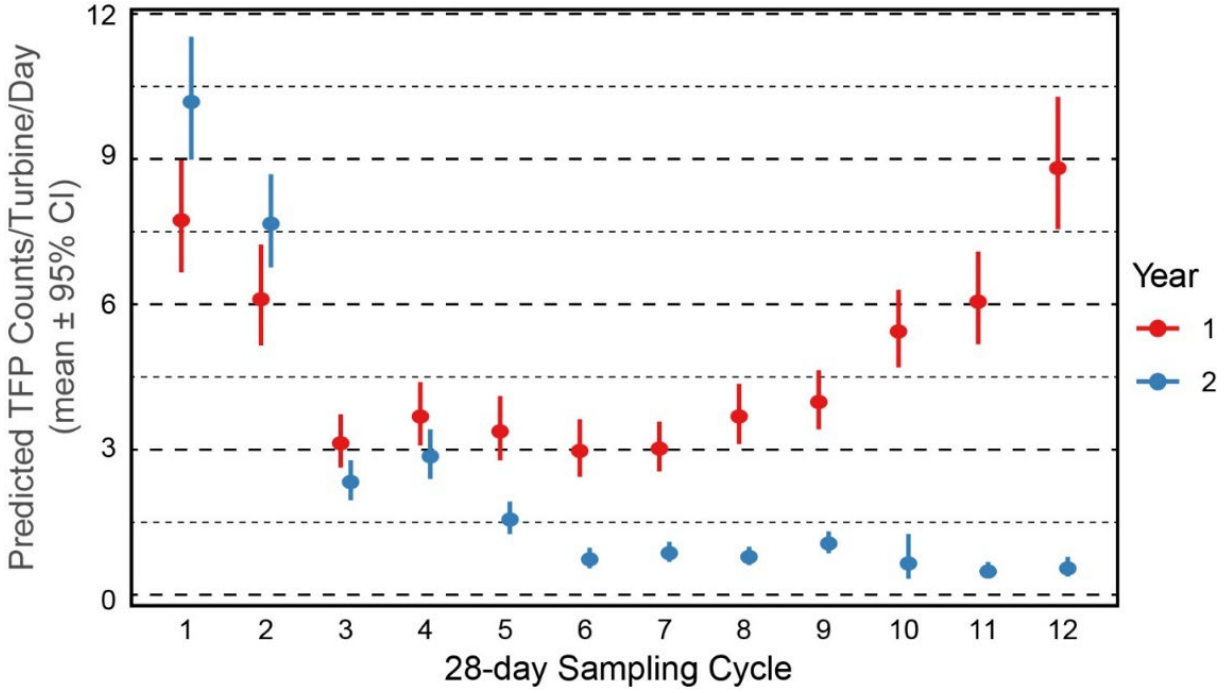

Figure 23. Predicted Average Daily Per Turbine True False Positive (TFP) DTBird Deterrent-Triggering Rates Across 28-day Sampling Periods During Study Years 1 and 2 at the Goodnoe Hills Wind Farm in Washington. Nonoverlapping Confidence Intervals Indicate Significant Pairwise Comparisons.

The GLMM relating daily turbine-specific counts of TFPs that triggered deterrents to *Site* and *28d Cycles* at the Manzanas and during Goodnoe Hills Year 2 revealed a highly significant main effect for *Site* (Wald χ^2, P <0.0001), a non-significant main effect for *28d Cycle* (P = 0.92), and a highly significant interaction term (P <0.0001). Nakagawa's marginal pseudo-R^2 for the model was 0.219, indicating the fixed effects provided moderate explanatory power. Planned post-hoc pairwise comparisons confirmed that (a) both sites had relatively elevated TFP rates during the first two *28d Cycles* of the respective sampling periods, (b) the early rates during Goodnoe Hills Year 2 were much higher than during the two corresponding cycles at the Manzanas, and (c) after adjustments were completed during the fifth cycle of Year 2 at the Goodnoe Hills, the TFP

deterrent-triggering rates followed similar patterns at the two sites, remained low and did not vary significantly across subsequent sampling cycles, and often were lower at the Goodnoe Hills post-adjustments than at the Manzanas (Figure 24).

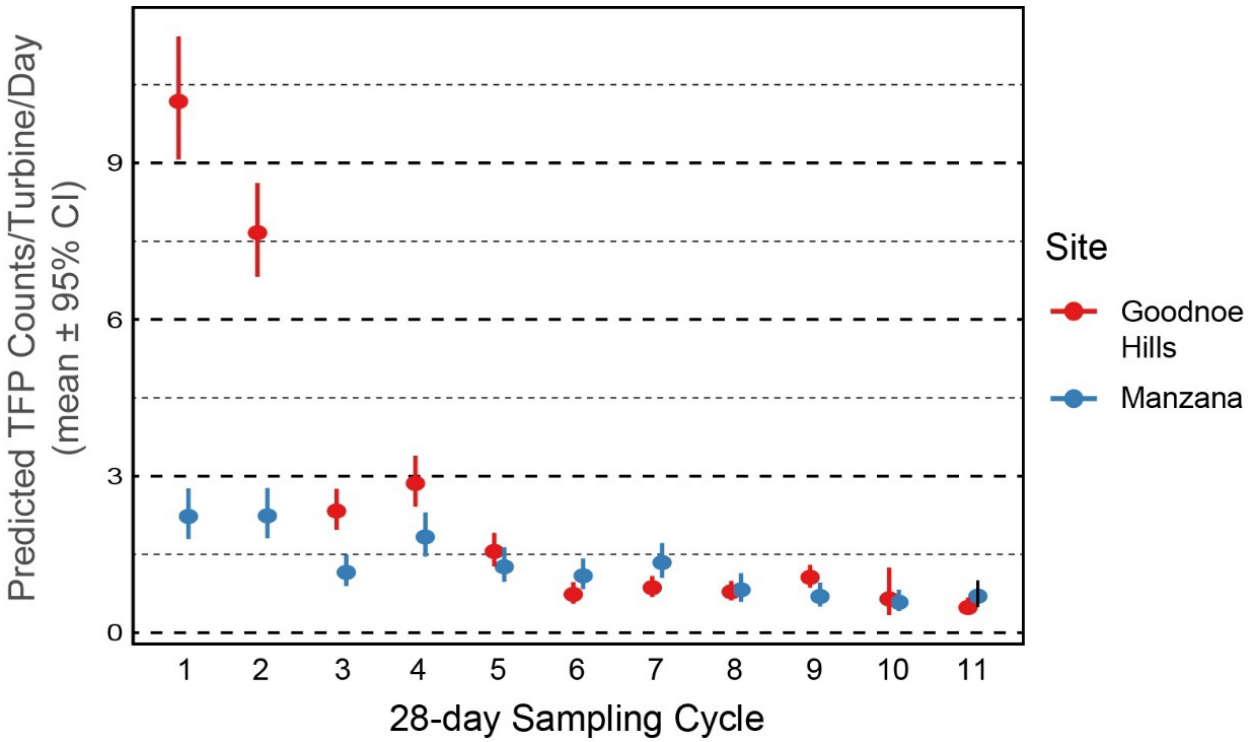

Figure 24. Predicted Average Daily Per Turbine True False Positive (TFP) DTBird Deterrent-Triggering Rates Across 12 28-day Sampling Periods (Variable Calendar Periods) at the Manzana Wind Power Project in California and During Study Year 2 at the Goodnoe Hills Wind Farm in Washington. Nonoverlapping Confidence Intervals Indicate Significant Pairwise Comparisons.

The GLMM relating daily turbine-specific counts of NTAFPs that triggered deterrents to *Year* and *Month* at the Goodnoe Hills revealed a highly significant main effect for *Year* (Wald χ^2, P <0.0001), a non-significant main effect for *28d Cycle* (P = 0.99), and a highly significant interaction term (P <0.0001). Nakagawa's marginal pseudo-R^2 for the model was 0.085, indicating the fixed effects provided marginal explanatory power. Unlike the TFP results, no dramatic shift in NTAFP prevalence occurred post-adjustments at the Goodnoe Hills; however, the post-adjustment rates in Year 2 (after January) did generally remain significantly lower than during all corresponding months in Year 1 (Figure 25).

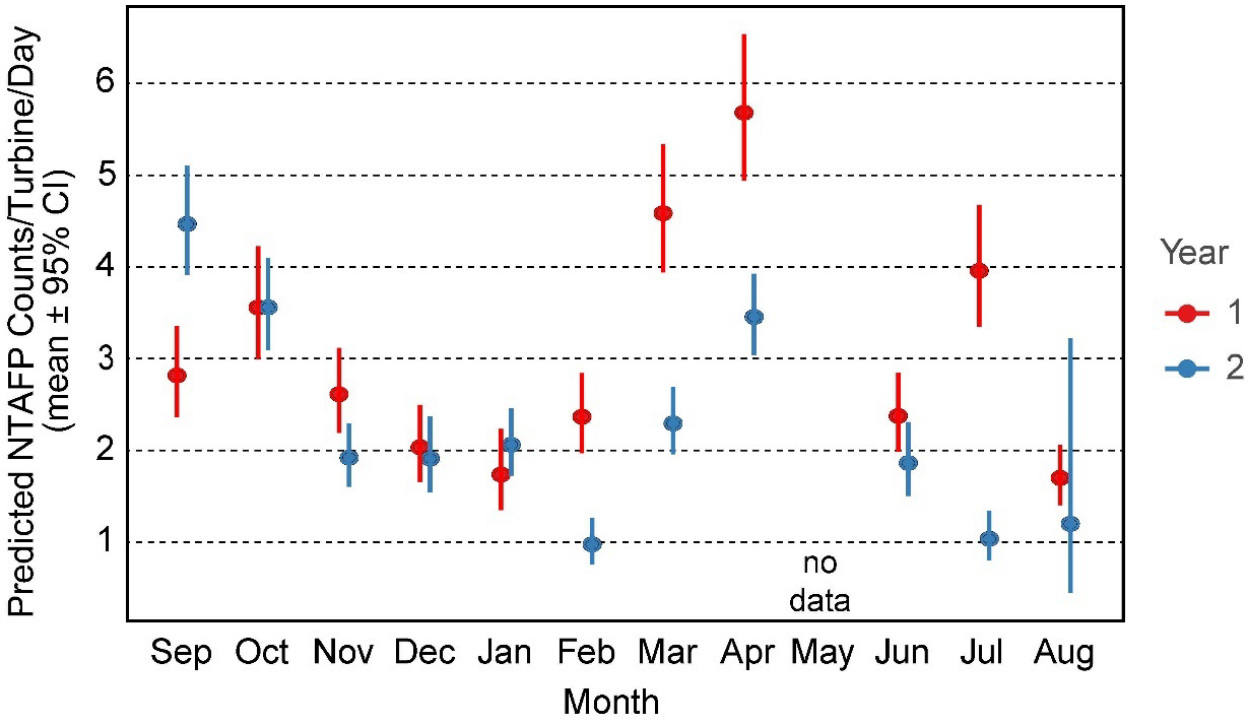

Figure 25. Predicted Average Daily Nontarget Avian False Positive (NTAFP) DTBird Deterrent-Triggering Rates Across 11 Months During Study Years 1 and 2 at the Goodnoe Hills Wind Farm in Washington. Nonoverlapping Confidence Intervals Indicate Significant Pairwise Comparisons.

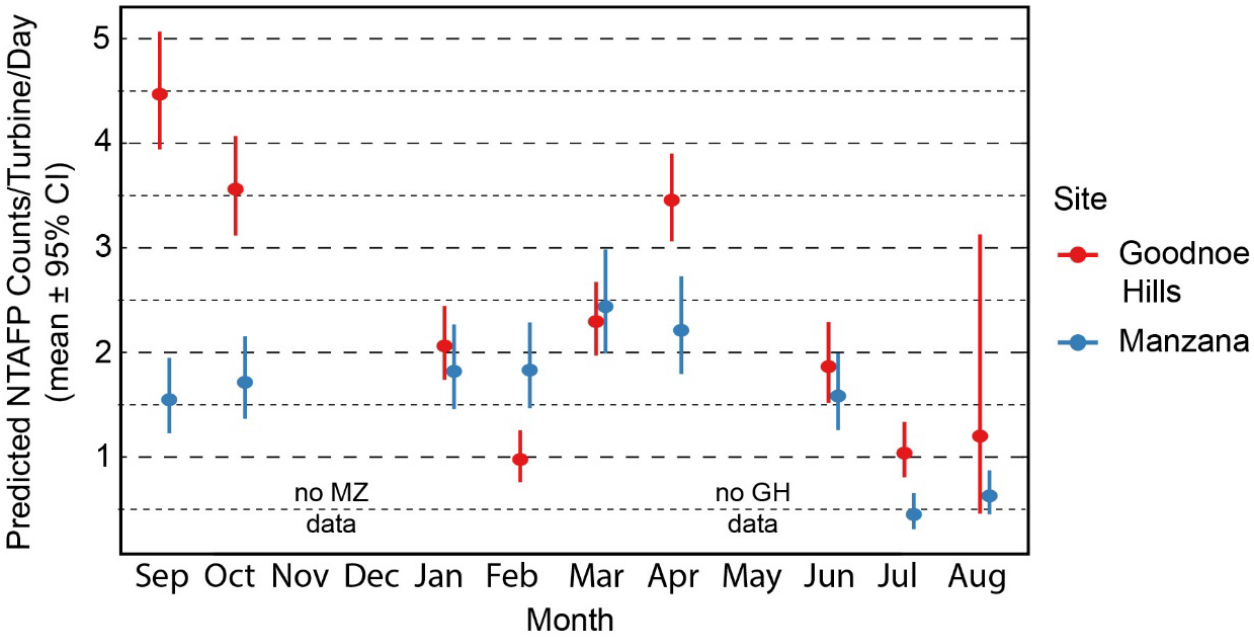

Figure 26. Predicted Average Daily Nontarget Avian False Positive (NTAFP) DTBird Deterrent-Triggering Rates Across Nine Common Sampling Months at the Manzana Wind Power Project in California and During Study Year 2 at the Goodnoe Hills Wind Farm in Washington. Nonoverlapping Confidence Intervals Indicate Significant Pairwise Comparisons.

The GLMM relating daily turbine-specific counts of NTAFPs that triggered deterrents to *Site* and *Month* at Manzana and at the Goodnoe Hills during sampling Year 2 revealed a non-significant main effect for *Site* (Wald χ^2, $P = 0.23$), a non-significant main effect for *28d Cycle* ($P = 0.98$), and a highly significant interaction term ($P < 0.0001$). Nakagawa's marginal pseudo-R^2 for the model was 0.129, indicating the fixed effects provided marginal explanatory power. Across the nine relevant calendar months, the two sites showed similarities towards higher NTAFP prevalence in spring, declining into mid-summer, then increasing some again in fall (Figure 26). The only substantive difference in pattern was that NTAFP prevalence was notably elevated at Goodnoe Hills during September and October compared to Manzana, suggesting higher fall migratory activity of nontarget birds at Goodnoe Hills.

4.5.4 In Situ Experimental Evaluation of Raptor Responses (Objective 4)

Table 22 summarizes the samples of confirmed and probable eagles we derived from screening DTBird event records on selected sample days, including the numbers of records for each species/group that did and did not trigger a deterrent signal under conditions when deterrent triggering was expected to occur if a bird passed within triggering range. These samples formed the basis for our analyses.

Table 2222. Summary of DTBird Detection Samples Used to Evaluate Results of Two-year Experiment Comparing Responses of Large Raptors to Muted (Control) Versus Broadcasted (Treatment) Audio Deterrents

Experiment Group – Species/Group[1]	Days With Samples	No Deterrence Records[2]	Deterrence Records[3]	Total Records	Average Records Per Day	SD
Control						
Golden Eagles	71	6	99	105	0.8	1.04
Bald Eagles	64	6	70	76	0.8	0.75
All Eagles	135	15	199	209	0.9	1.18
Treatment						
Golden Eagles	70	11	91	102	0.8	1.11
Bald Eagles	40	2	51	53	0.5	0.72
All Eagles	123	13	168	181	0.8	1.05

[1] In all cases, classifications include confirmed and probable identifications belonging to the specific species or species group.

[2] Cases where a target bird was detected but did not trigger a deterrent signal.

[3] Cases where a target bird was detected and triggered one or both deterrent signals, either virtually (control mode) or with the deterrents actually broadcasting (treatment mode).

4.5.4.1 Testing Hypothesis A Regarding Probability of Eagles Triggering a DTBird Dissuasion Signal

For confirmed and probable golden eagles alone, limited sample sizes constrained our ability to evaluate a full model including the complete suite of potential predictors and 2-way interactions of interest. Instead, we proceeded systematically to evaluate (1) the influences of *Treatment Group* combined with each of the other predictors alone and then with associated two-way interactions, and (2) more complex multi-variable models based on indications of potential significance during the preceding step (see Attachment 3: Appendix C for comparisons of selected candidate models). Throughout the process of considering candidate models and selecting a final logistic GLMM to represent the probability of golden eagles triggering a dissuasion signal, the prediction coefficients for *Treatment Group* were always negative, suggesting the expected effect of a lower probability of dissuasion triggers at turbines operating in treatment mode. *Treatment Group* never emerged as even a marginally significant predictor, however. In contrast, *Year*, *Time of Day*, and *Wind Speed* were at least marginally significant predictors and were retained in the final model. Accordingly, the dissuasion-trigger model selected to represent golden eagles alone, based on AIC scores, parameter tests, and positive model diagnostics, was as follows:

> ln(Odds of dissuasion trigger) ~ [1|*Turbine ID*] + [1|*Turbine ID* : *Elapsed Days*] + *Treatment Group* + *Year* + *Time of Day* + *Wind Speed*

The relationships indicated by the resulting model coefficients and individual parameter tests (Table 23) are described below.

- Non-significant 29% reduction (95% CI: 63% reduction – 36% increase) in the probability of dissuasion triggers at installations operating in treatment mode.
- Marginally significant 46% reduction (95% CI: 73% reduction – 9% increase) in the probability of dissuasion triggers in *Year 2*.

- Marginally significant positive relationship between the probability of dissuasion triggers and *Time of Day* (Figure 27).
- Significant negative relationship between the probability of dissuasion triggers and *Wind Speed* (Figure 28).

Table 2323. Model Coefficients and Fixed Effect Parameter Test Results for the Logistic GLMM Selected to Represent the Probability of Confirmed and Probable Golden Eagles Triggering a Dissuasion Signal at DTBird Installations Operating in Treatment (Deterrents Broadcasting) and Control (Deterrents Muted) Mode During Two-year Experiment

Random Effect	Variance	SD				
Turbine	0.357	0.5977				
Turbine: Elapsed Days[1]	0.116	0.3409				
Fixed Effect	**Estimate**	**SE**	***z*** [2]	***P* (>|z|)** [2]	**LRT χ^2** [3]	***P* (>χ^2)** [3]
Intercept	0.546	0.3421	1.597	0.110	–	–
Treatment Group: On[4]	-0.339	0.3304	-1.026	0.305	1.07	0.302
Year: 2 [5]	-0.614	0.3569	-1.721	0.085	3.06	0.080
Time of Day[6]	0.295	0.154	1.917	0.055	3.83	0.050
Wind Speed[7]	-0.385	0.1780	-2.161	0.031	5.10	0.024

[1] Elapsed Days = days since data-collection began; a simpler equivalent of date.

[2] Wald test.

[3] Drop1 likelihood ratio test.

[4] Reference category – Off = control mode. On = treatment mode.

[5] Reference category – Year 1: 1 September 2021 – 31 August 2022. Year 2: 1 September 2022 – 30 September 2023 (extended due to facility shut down from 1–24 May 2023.

[6] Translated to minutes of the day; centered and scaled ([value – mean]/SD).

[7] Recorded in meters/second; centered and scaled ([value – mean]/SD).

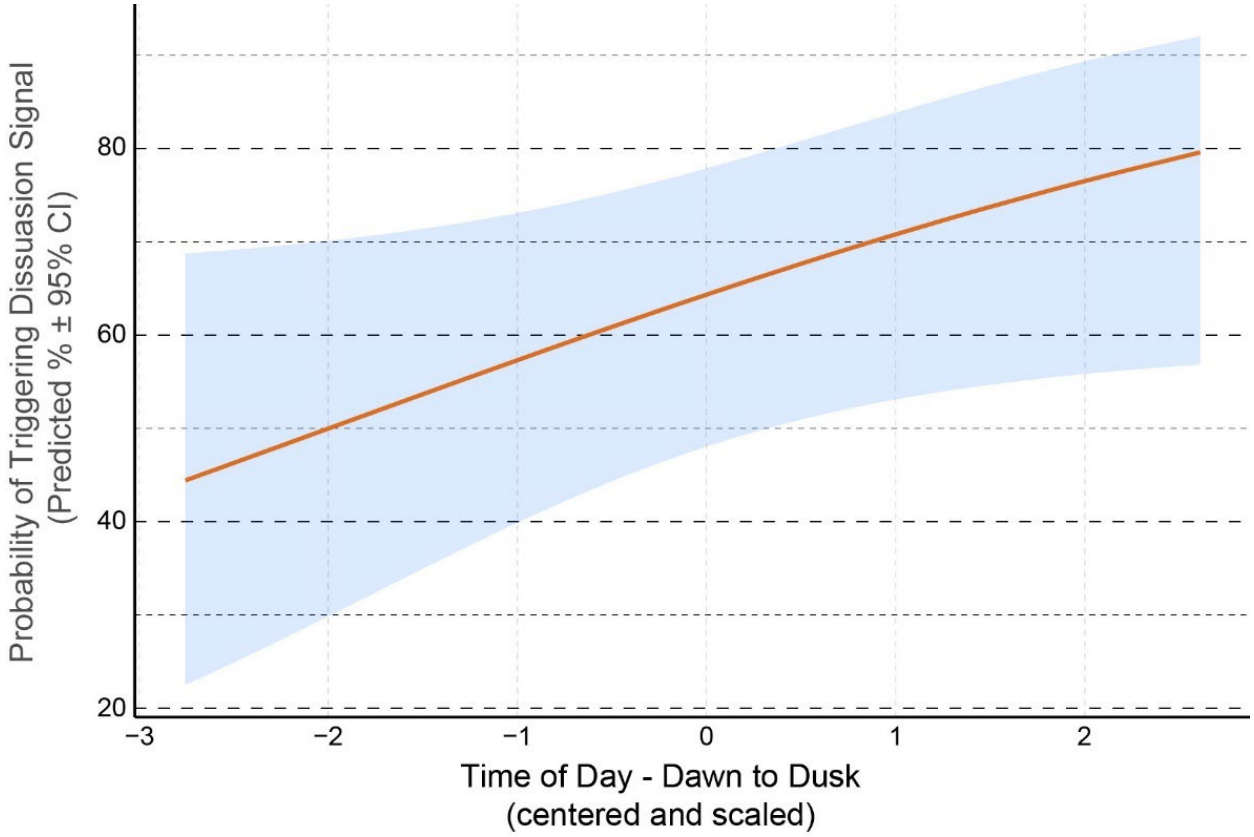

Figure 27. Illustration of predicted relationship between the probability of a golden eagle triggering a DTBird dissuasion signal and *Time of Day*.

Based on the dataset limited to eagles positively identified as either a golden eagle or a bald eagle, again no significant *Treatment Group* effects were evident but other indicators similar to

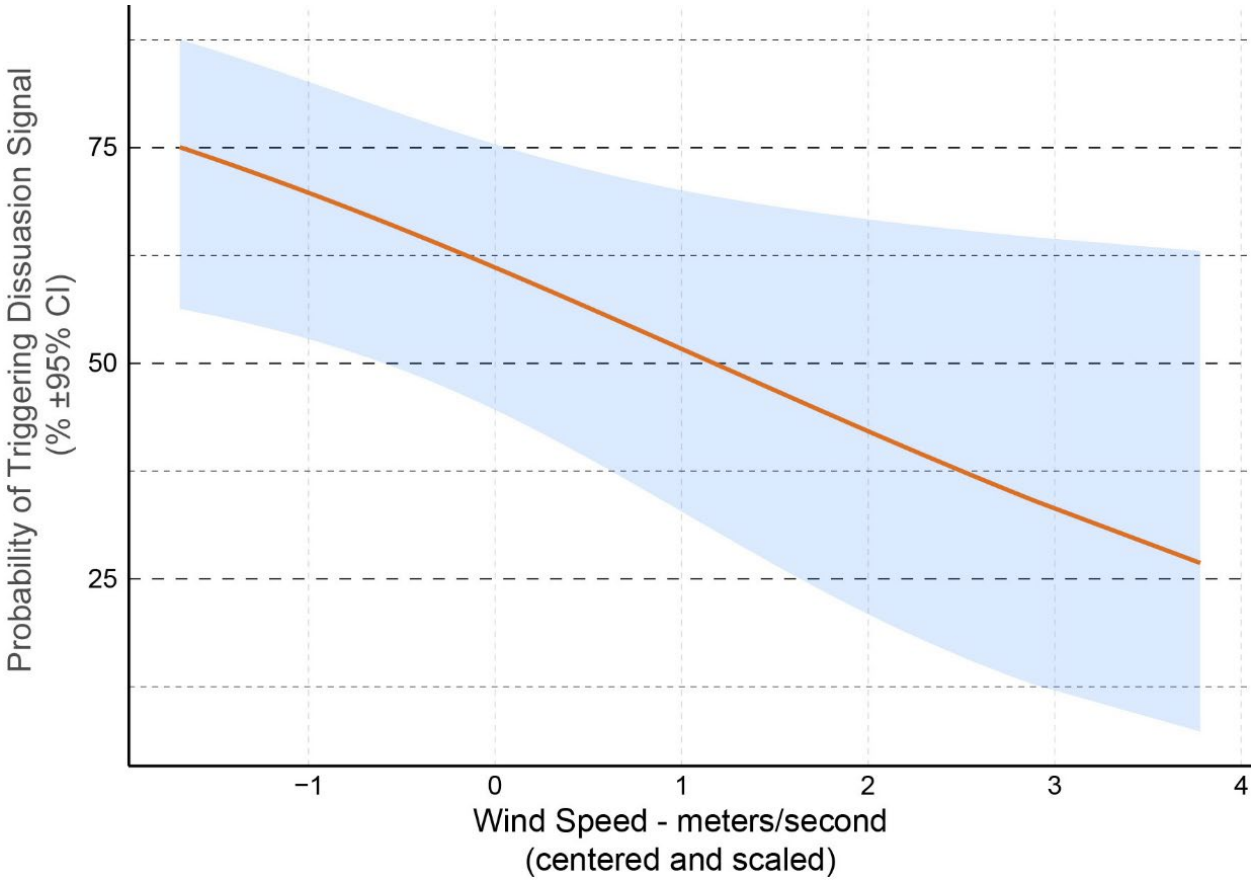

Figure 28. Illustration of predicted relationship between the probability of a golden eagle triggering a DTBird dissuasion signal and *Wind Speed*.

the results for golden eagles alone were evident. More importantly, although preliminary indications emerged suggesting potential marginal differences in the probability of dissuasion triggering for the two eagle species, those indications faded away once other covariates were included in the models. Therefore, we abandoned further consideration of models limited to identified golden and bald eagles with *Species* as a predictor in favor of evaluating models based on the larger all-eagles dataset (see Table 22) without considering *Species* as a potential predictor. Based on this dataset, we were able take both full backwards and forwards stepwise model building approaches to identify a top model (see Attachment 3: Appendix D for comparisons of models evaluated as part of a backwards elimination process to select the final model). The outcome of this approach again did not reveal a strong *Treatment Group* effect; however, the selected model included two at least marginally significant interactions between *Treatment Group* and other predictors, which provided important insight. The structure of the dissuasion-trigger logistic GLMM selected to represent all eagles combined was as follows:

> ln(Odds of dissuasion trigger) ~ [1|*Turbine ID*] + [1|*Turbine ID* : *Elapsed Days*] + *Treatment Group* + *Time of Day* + *Time of Day*^{2} +*Cloud Cover* + *FPs per Day* + *Treatment Group* * *Cloud Cover* + *Treatment Group* * *FPs per Day*

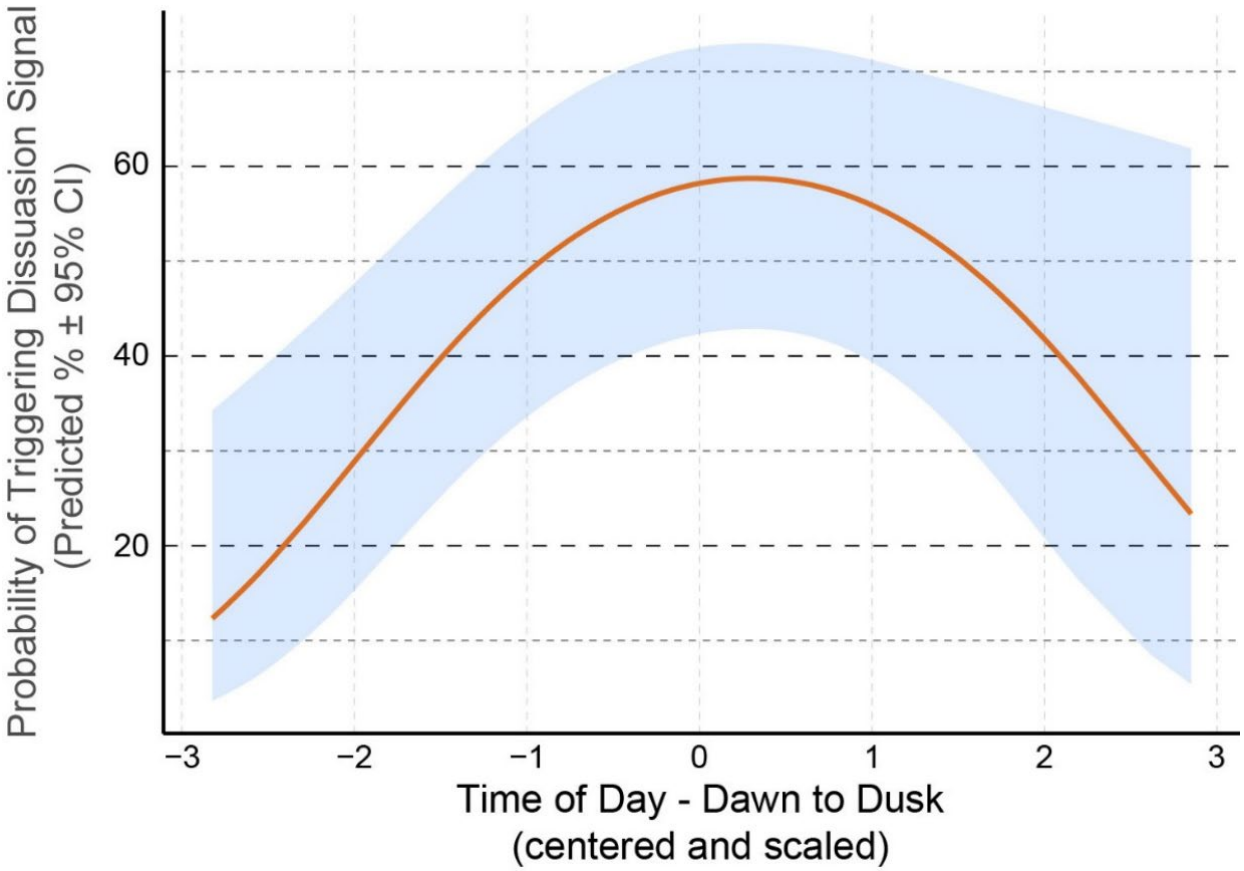

Figure 29. Illustration of predicted second-order relationship between the probability of an eagle triggering a DTBird dissuasion signal and time of day.

The relationships indicated by the resulting model coefficients and individual parameter tests (Table 24) are described below.

- Significant second-order relationship between the probability of dissuasion triggers and *Time of Day*, reflecting a higher probability of dissuasion triggering during midday compared to earlier and later in the day (Figure 29).
- When partly cloudy, cloudy, or overcast skies prevailed, the probability of dissuasion triggers was at least slightly lower at turbines operating in treatment mode compared to those operating in control mode, whereas when fair skies prevailed, the probability of dissuasion triggers was substantially lower at turbines operating in control mode (Figure 30).
- At turbines with DTBird systems operating in control mode, the probability of dissuasion triggers increased as the number of *FPs per Day* increased, whereas the opposite pattern applied at turbines operating in treatment mode (Figure 31).

Table 2424. Model Coefficients and Fixed Effect Parameter Test Results for the Logistic GLMM Selected to Represent the Probability of Confirmed and Probable Eagles (Golden and Bald Eagles Combined) Triggering a Dissuasion Signal at DTBird Installations Operating in Treatment (Deterrents Broadcasting) and Control (Deterrents Muted) Mode During Two-year Experiment

Random Effect	Variance	SD				
Turbine ID	0.285	0.5338				
Turbine ID: Elapsed Days[1]	0.389	0.624				
Fixed Effect	**Estimate**	**SE**	***z*** [2]	***P* (>|z|)** [2]	**LRT χ^2** [3]	***P* (>χ^2)** [3]
Intercept	0.374	0.3292	1.136	0.256	–	–
Treatment Group: On[4]	-0.263	0.3911	-0.672	0.501	–	–
Cloud Cover: Fair[5]	-1.278	0.5399	-2.367	0.018	–	–
Cloud Cover: Overcast[5]	0.377	0.5757	0.655	0.512	–	–
Cloud Cover: Partly Cloudy[5]	1.133	0.4120	2.751	0.006	–	–
Time of Day[6]	0.143	0.1226	1.165	0.244	1.359	0.244
Time of Day[6]	-0.237	0.0888	-2.668	0.008	7.939	0.004
FPs per Day[7]	0.395	0.1802	2.192	0.028	–	–
Treatment Group * Cloud Cover: Fair	2.040	0.7363	2.771	0.006	16.254	0.001
Treatment Group * Cloud Cover: Overcast	-0.297	0.8010	-0.371	0.710	–	–
Treatment Group * Cloud Cover: Partly Cloudy	-0.909	0.6004	-1.514	0.130	–	–
Treatment Group * FPs per Day	-0.492	0.2811	-1.750	0.080	2.965	0.085

[1] Elapsed Days = days since data-collection began; a simpler equivalent of date.

[2] Wald test.

[3] Drop1 likelihood ratio test.

[4] Reference category – Off = control mode. On = treatment mode.

[5] Reference category – Cloudy. Fair = mostly cloud free; Partly cloudy = <50% cloud cover; Cloudy = ≥50% cloud cover with distinctly variable cloud definitions and brightness; Overcast = complete and largely uniform gray or darker cloud cover.

[6] Translated to minutes of the day; centered and scaled ([value – mean]/SD).

[7] FPs = false positives. Number of detection events triggered by true FPs and non-target avian FPs (see Section 2.4); centered and scaled ([value – mean]/SD).

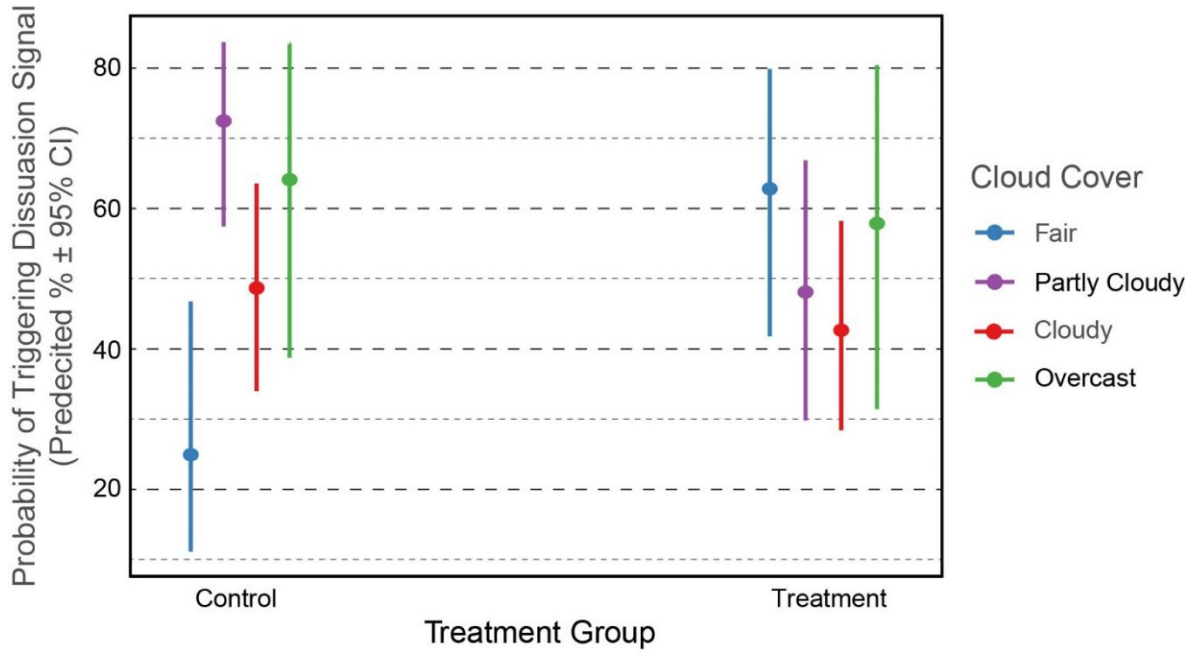

Figure 30. Illustration of predicted interactive relationship between treatment group and cloud cover in determining the probability of an eagle triggering a DTBird dissuasion signal.

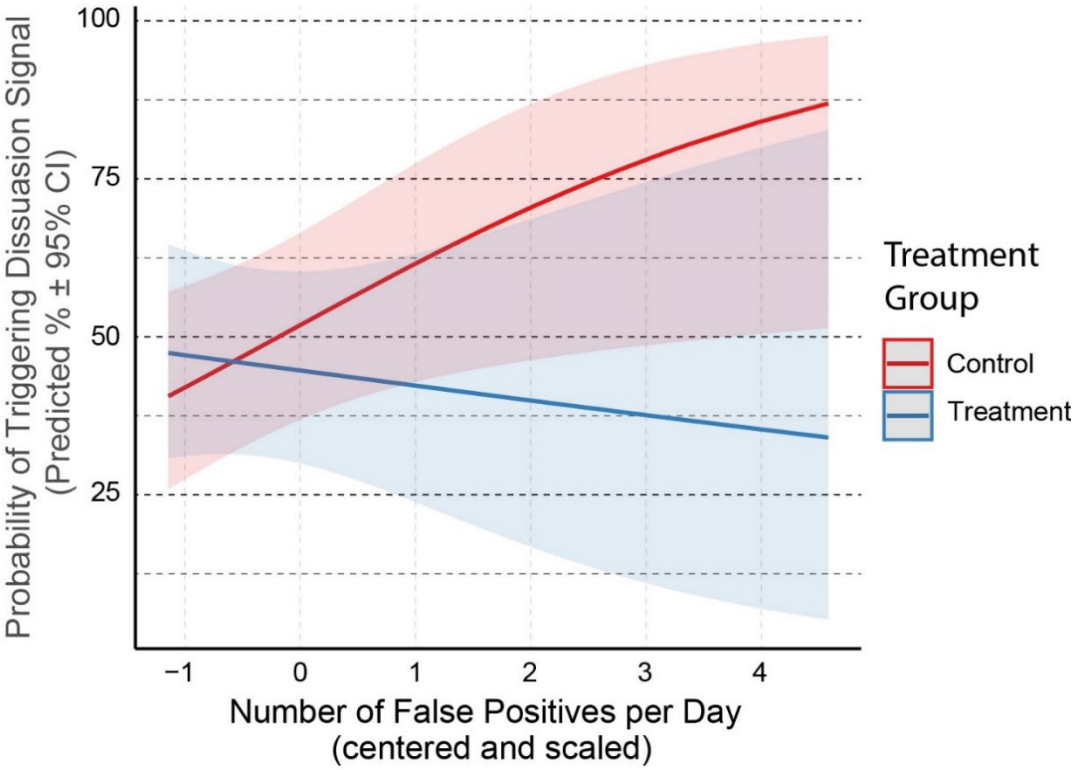

Figure 31. Illustration of predicted interactive relationship between treatment group and the daily numbers of false positives that triggered deterrent signals in determining the probability of an eagle (golden and bald eagles combined) triggering a DTBird dissuasion signal.

Augmenting the selected model above by including *Wind Speed* resulted in the lowest AIC score among the evaluated models (Attachment 3: Appendix D); however, the ΔAIC was only 0.4 points and the *P* value for the likelihood ratio test evaluating the contribution of *Wind Speed* to the model (0.118) did not meet our criterion for retention in the model. Nevertheless, the negative parameter coefficient indicated a similar pattern as the significant relationship indicated for golden eagles alone, suggesting that wind speeds might have differentially influenced the responses of golden and bald eagles around the Goodnoe Hills turbines.

4.5.4.2 Testing Hypothesis B Regarding Dwell Time of Eagles Around DTBird Deterrent Systems

To develop the GLMM for evaluating the influence of *Treatment Group* and other potential predictors on the dwell time of golden eagles around the study turbines, we were able take both full backwards and forwards stepwise model building approaches to identify a top model. The resulting selected model (see Attachment 3: Appendix E for comparisons of models evaluated as part of a backwards elimination process to select the final model) had the following form:

> Dwell Time ~ [1|*Turbine ID*] + [1|*Turbine ID* : *Elapsed Days*] + *Treatment Group* + *28d Cycle* + *Time of Day* + *Time of Day2* + *FPs per Day* + *Treatment Group * FPs per Day*

The relationships indicated by the resulting model coefficients and individual parameter tests (Table 25) are described below.

- Significant 27% reduction (95% CI: 5–42%) in the average dwell time of golden eagles at installations operating in treatment mode, with the average dwell time reduced from approximately 26 to 17 seconds per event.
- Marginally significant overall declining trend in the dwell time of golden eagles in relation to the progression of *28d Cycles* over the course of the two-year study (Figure 32).
- Significant main effect / marginally significant second-order relationship between dwell time and *Time of Day*, reflecting short dwell times in the morning, increasing through mid-afternoon, then tapering off again in the evening (Figure 33).
- Marginally significant interaction between *Treatment Group* and *FPs per Day* illustrating a positive relationship between dwell times and FP numbers around control turbines, but a negative relationship around treatment turbines (Figure 34). Put another way, the more that FPs contributed to actual deterrent broadcasting at treatment turbines, the less likely were eagles to dwell in the vicinity of those turbines.

Table 2525. Model Coefficients and Fixed Effect Parameter Test Results for the GLMM Selected to Represent the Relationship Between the Dwell Time of Confirmed and Probable Golden Eagles at DTBird Installations Operating in Treatment (Deterrents Broadcasting) and Control (Deterrents Muted) Mode During Two-year Experiment.

Random Effect	Variance	SD
Turbine	0.014	0.1166
Turbine: Elapsed Days[1]	2.15E-07	0.0005

Fixed Effect	Estimate	SE	*z* [2]	*P* (>|z|) [2]	LRT χ^2 [3]	*P* (>χ^2) [3]
Intercept	3.304	0.1082	30.54	<0.001	–	–
Treatment Group: On[4]	-0.319	0.1258	-2.54	0.011	–	–
28d Cycle[5]	-0.135	0.0661	-2.04	0.041	4.08	0.044
Time of Day[6]	0.166	0.0666	2.50	0.013	6.42	0.011
Time of Day^2	-0.089	0.0451	-1.98	0.047	3.66	0.056
FPs per Day[7]	0.086	0.0754	1.14	0.255	–	–
Treatment Group: On * FPs per Day	-0.258	0.1361	-1.90	0.058	3.22	0.073

[1] Elapsed Days = days since data-collection began; a simpler equivalent of date.
[2] Wald test.
[3] Drop1 likelihood ratio test.
[4] Reference category – Off = control mode. On = treatment mode.
[5] Discrete continuous predictor representing 27 consecutive 28-day sampling periods from 1 September 2021 through 30 September 2023.
[6] Translated to minutes of the day; centered and scaled ([value – mean]/SD).
[7] FPs = false positives. Number of detection events triggered by true FPs and non-target avian FPs (see Section 2.4); centered and scaled ([value – mean]/SD).

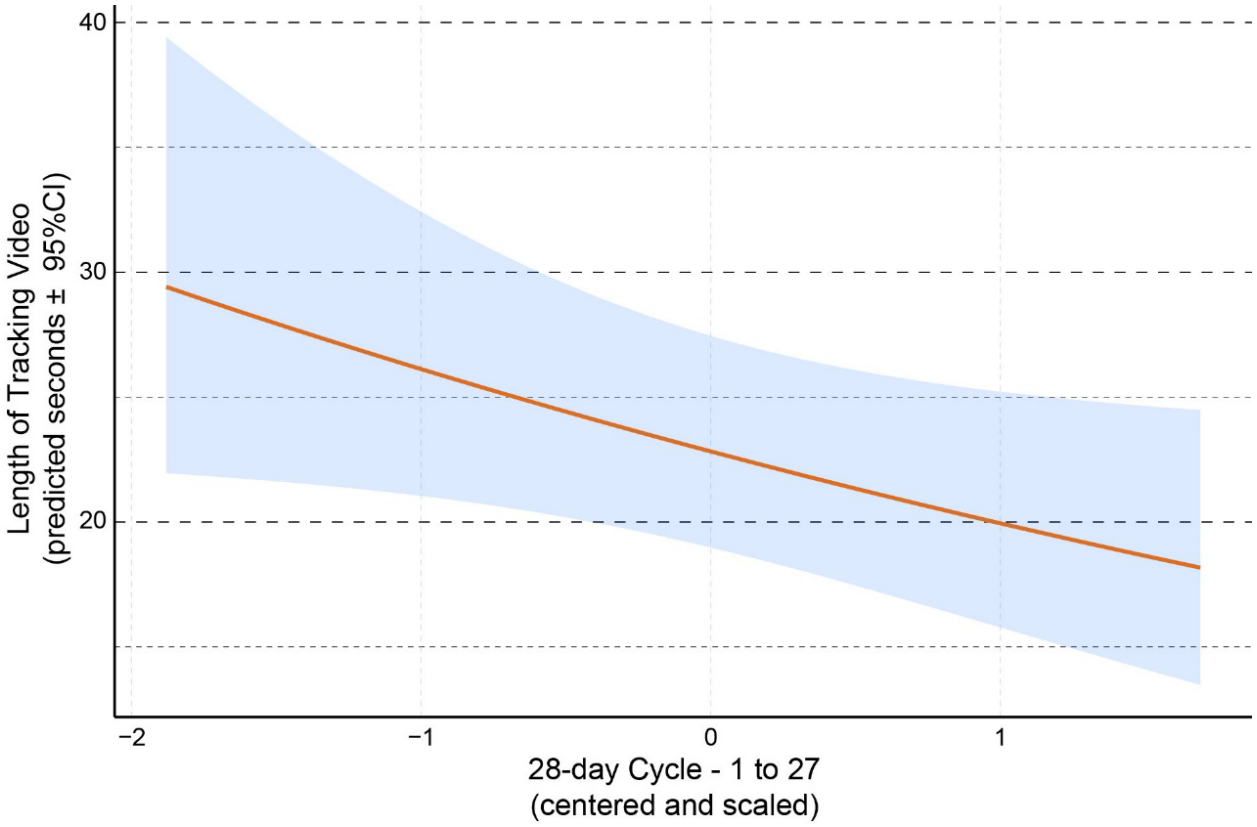

Figure 32. Illustration of predicted decline in the dwell time of golden eagles at DTBird turbines across the 27 28-day sampling cycles that composed this two-year experimental analysis.

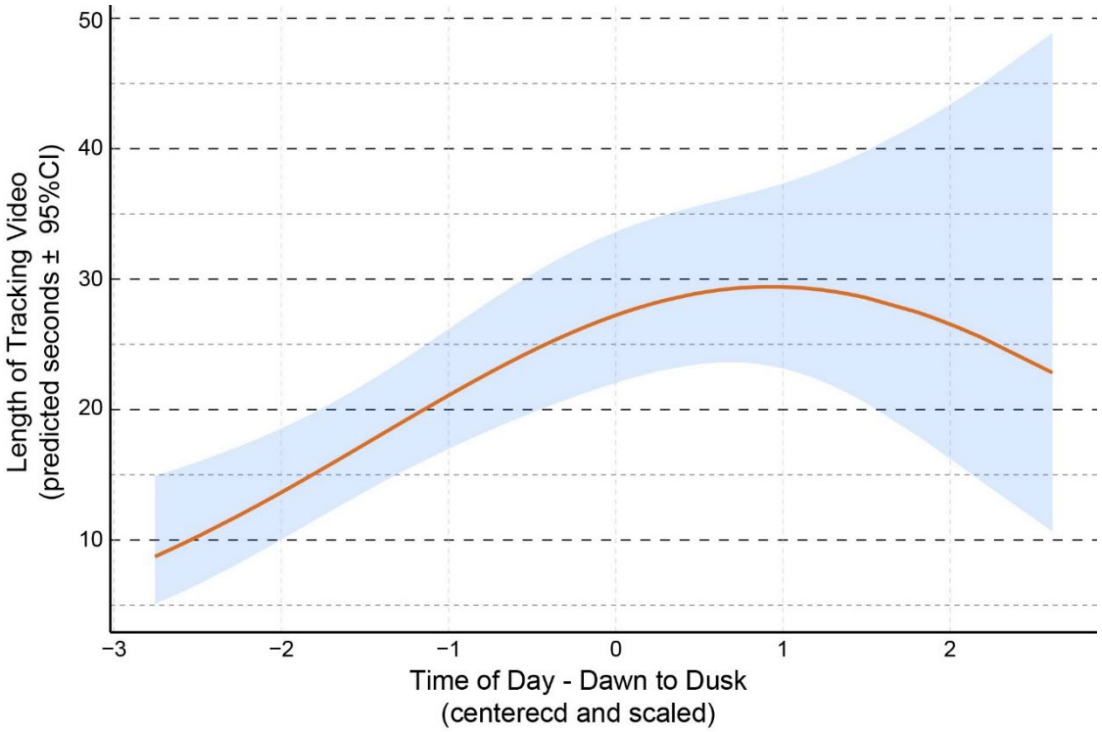

Figure 34. Illustration of predicted second-order relationship between the dwell time of golden eagles at DTBird turbines and time of day.

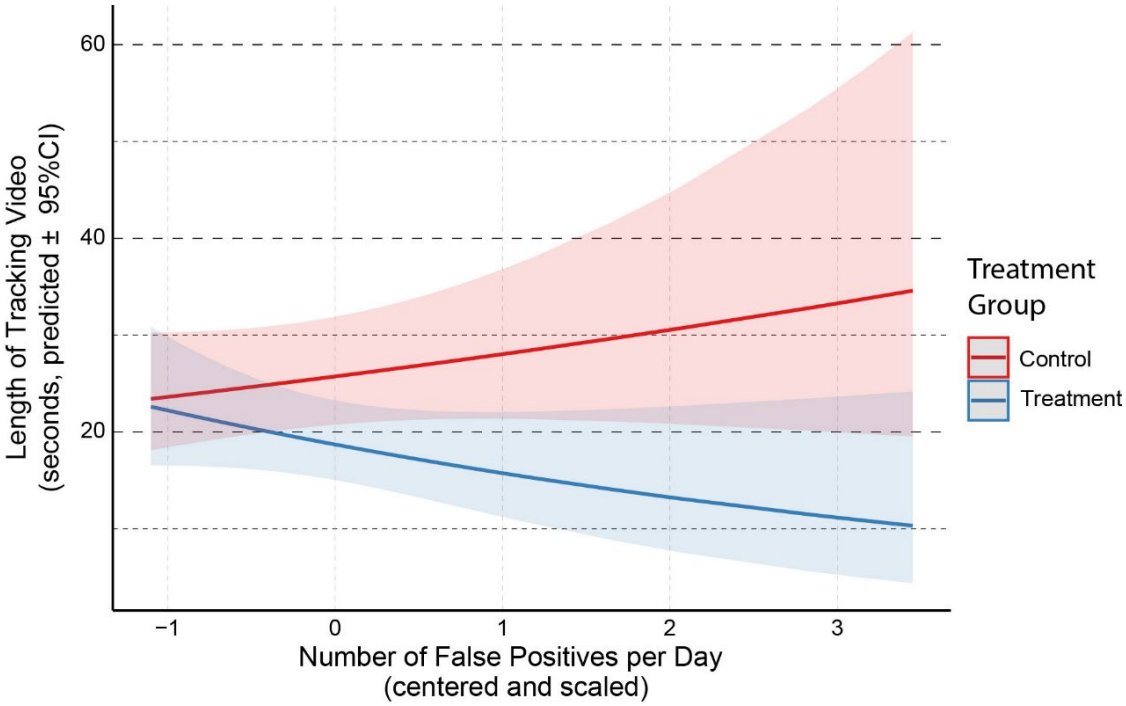

Figure 33. Illustration of predicted interactive relationship between treatment group and the daily numbers of false positives (FPs) that triggered deterrent signals in determining the dwell time of golden eagles around DTBird turbines.

Considering the dataset limited to eagles positively identified as either a golden eagle or a bald eagle yielded no evidence of *Species* as an influential predictor of dwell time. Hence, again we focused our further attention on evaluating models based on the larger all-eagles dataset without considering *Species* as a potential predictor. Running full models based on this dataset and dependent variable proved untenable due to dataset limitations; hence, we proceeded to identify a top model based on a similar iterative approach as described for golden eagles alone. The outcomes of this modeling effort yielded similar insights as for predicting the dwell time of golden eagles alone, with the same final model selected to represent all eagles combined (see Attachment 3: Appendix F for comparisons of selected candidate models) and the model coefficients confirming similar relationships as described above (Table 25, Figures 32-34). Most germane was a significant estimated 24% reduction (95% CI: 7–35%) in the dwell time of eagles at treatment turbines, with the average dwell time reduced from approximately 25 to 19 seconds per event. Note that, in deciding upon a final dwell-time model for all eagles combined, we retained *FPs per Day* and the *Treatment Group * FPs per Day* interaction (see Figure 34) despite the *P* value for the interaction (0.129) being slightly greater than our *P* ≤0.10 threshold for inclusion. We did this to retain a relationship that improved the AIC score of the final model and was common to two of the other three primary models we evaluated—albeit only marginally significant in each case (see Tables 25 and 26).

Table 2626. Model Coefficients and Fixed Effect Parameter Test Results for the GLMM Selected to Represent the Relationship Between the Dwell Time of All Confirmed and Probable Eagles at DTBird Installations Operating in Treatment (Deterrents Broadcasting) and Control (Deterrents Muted) Mode During Two-year Experiment.

Random Effect	Variance	SD				
Turbine	0.0016	0.03406				
Turbine: Elapsed Days[1]	1.82E-08	0.00014				
Fixed Effect	**Estimate**	**SE**	***z*** [2]	***P* (>|z|)** [2]	**LRT χ^2** [3]	***P* (>χ^2)** [3]
Intercept	3.305	0.0729	45.33	<0.001	–	–
Treatment Group: On[4]	-0.269	0.0934	-2.88	0.004	–	–
28d Cycle[5]	-0.114	0.0479	-2.37	0.018	5.64	0.018
Time of Day[6]	0.093	0.0453	2.09	0.037	4.49	0.034
Time of Day^2	-0.093	0.0316	-2.93	0.003	7.92	0.005
FPs per Day[7]	0.124	0.0557	2.23	0.026	–	–
Treatment Group: On * FPs per Day	-0.149	0.0964	-1.55	0.121	2.31	0.129

[1] Elapsed Days = days since data-collection began; a simpler equivalent of date.

[2] Wald test.

[3] Drop1 likelihood ratio test.

[4] Reference category – Off = control mode. On = treatment mode.

[5] Discrete continuous predictor representing 27 consecutive 28-day sampling periods from 1 September 2021 through 30 September 2023.

[6] Translated to minutes of the day; centered and scaled ([value – mean]/SD).

[7] FPs = false positives. Number of detection events triggered by true FPs and non-target avian FPs (see Section 2.4); centered and scaled ([value – mean]/SD).

4.5.4.3 Testing Hypothesis C Regarding the Probability of Eagles Crossing the Rotor Swept Area of DTBird Equipped Turbines

Modeling the probability of an RSA crossing for golden eagles alone and for all eagles combined yielded no *Treatment Group* effects and no models that improved upon the null model. This outcome was not surprising given a paucity of consistent and reliable data to evaluate this dependent variable. Observations recorded by our data-entry technicians suggested that 9% of 105 golden eagle observations at turbines with DTBird systems operating in control mode a potential RSA cross, whereas a nominally lower 7% of 102 observations at turbines operating in treatment mode included a potential RSA cross. For all eagles combined, the comparisons were 13% of 209 observations included a potential RSA cross at control turbines, and 12% of 181 observations included a potential RSA cross at treatment turbines.

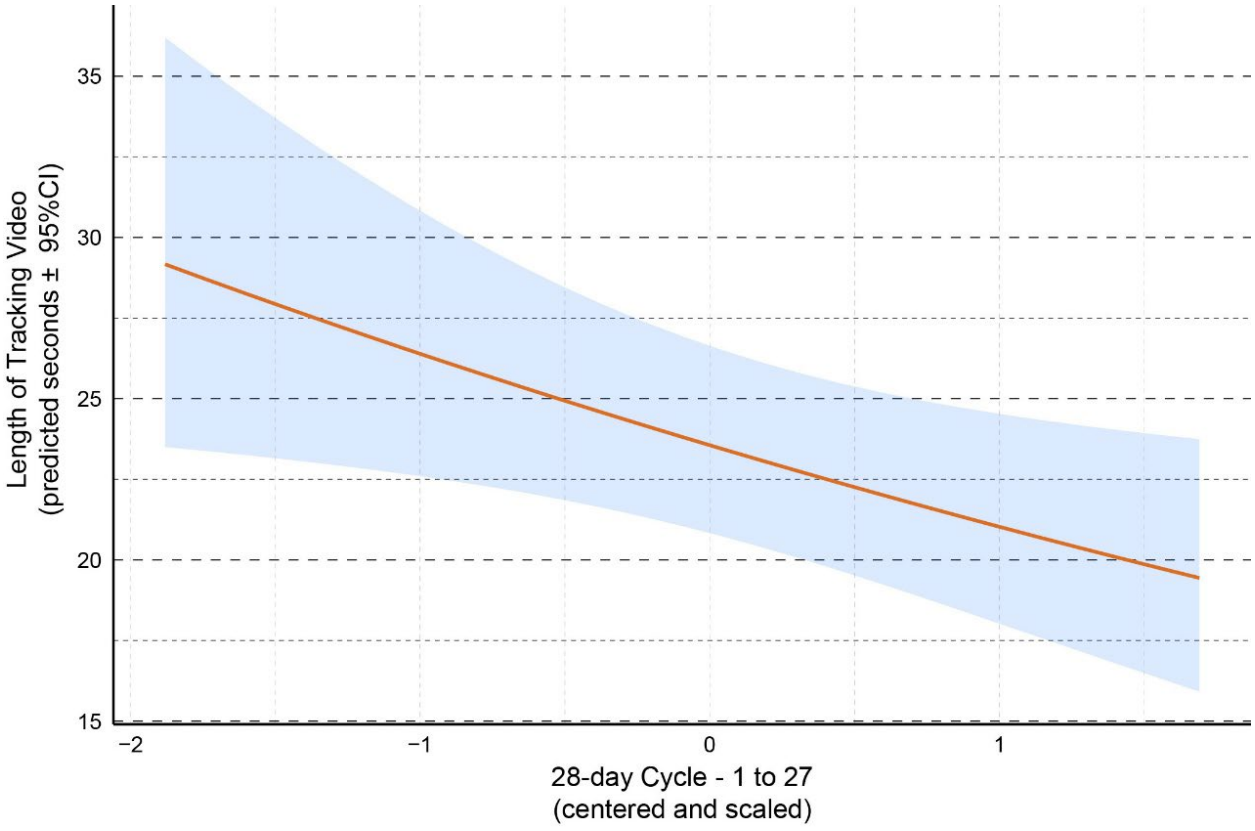

Figure 35. Illustration of predicted decline in the dwell time of eagles (golden and bald eagles combined) at DTBird turbines across the 27 28-day sampling cycles that composed this two-year experimental analysis.

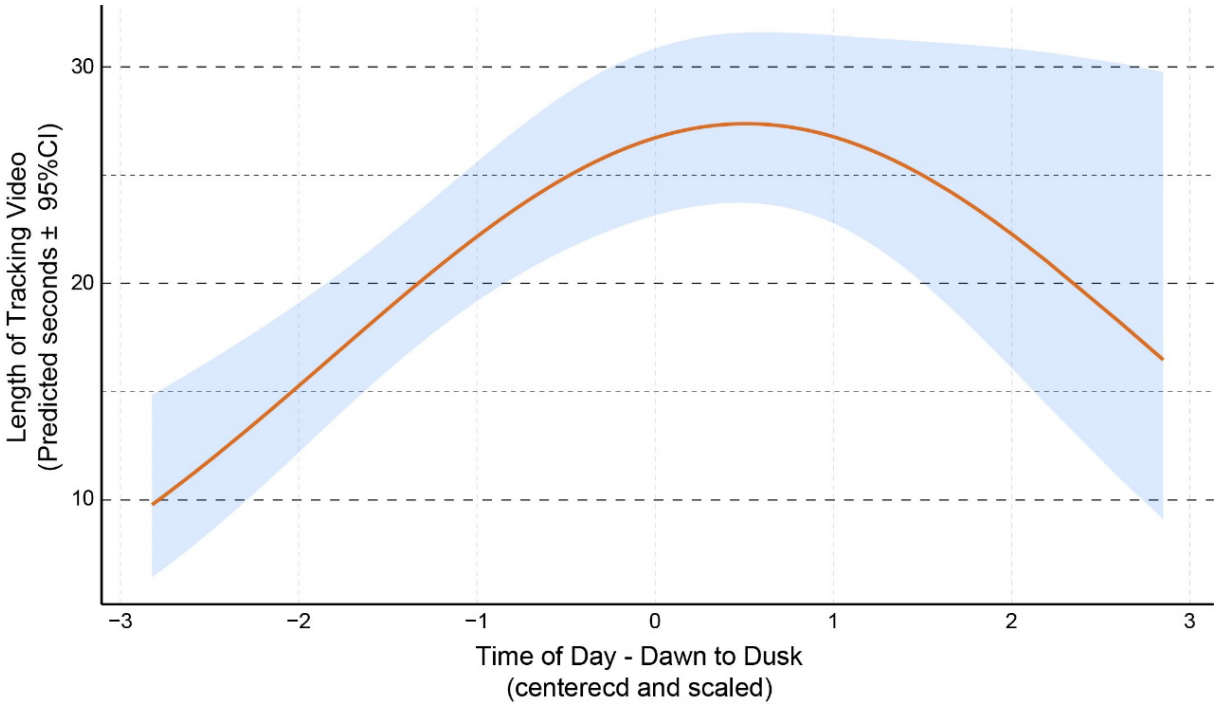

Figure 37. Illustration of predicted second-order relationship between the dwell time of eagles (golden and bald eagles combined) at DTBird turbines and time of day.

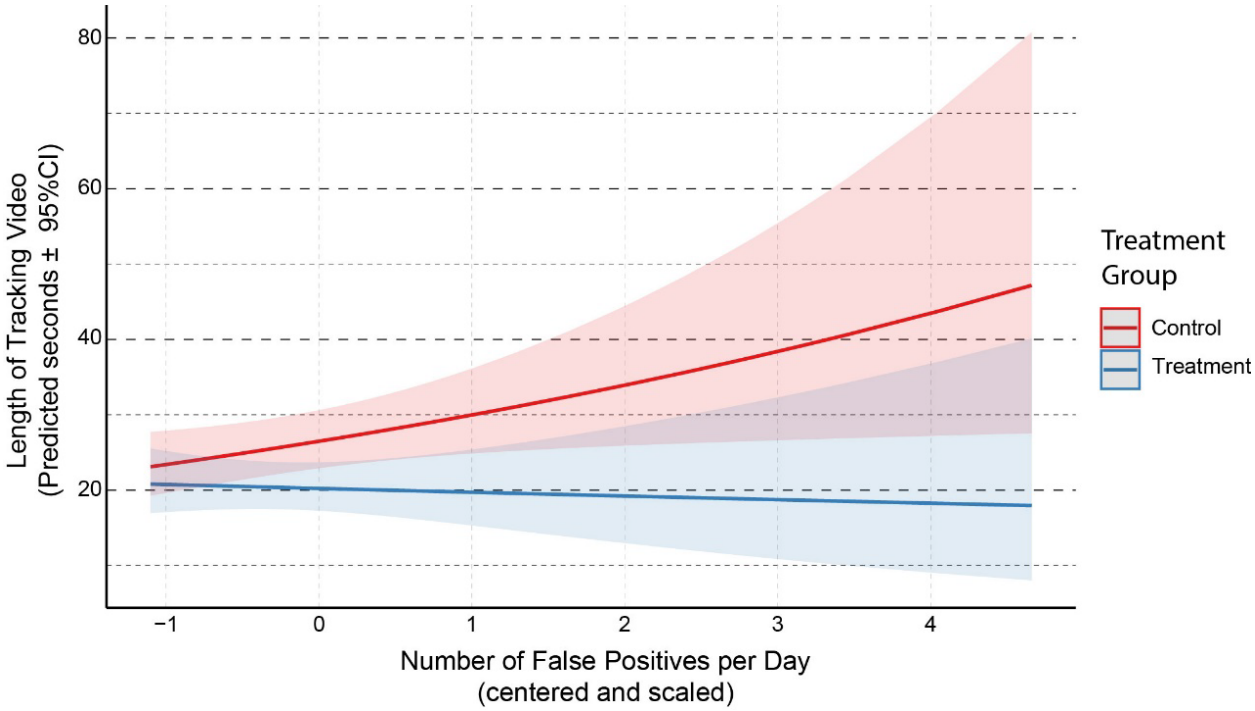

Figure 36. Illustration of predicted interactive relationship between treatment group and the daily numbers of false positives (FPs) that triggered deterrent signals in determining the dwell time of eagles (golden and bald eagles combined) around DTBird turbines.

4.5.5 Multi-site Analysis of Collision Risk Reduction (Objective 5)

Our first approach to estimating the overall effectiveness of DTBird in reducing the risk of eagles entering the RSZ of spinning turbines involves the product of the estimated overall probability of detection from the UAV flight trials and the estimated probability of presumed effective deterrence from the behavioral analysis. For golden eagles alone, the results suggested variable performance at the two study sites as follows:

> Manzana: 66% probability of detection x 79% probability of effective deterrence = 52% probability of reducing risk of entering RSZ of spinning turbines
>
> Goodnoe Hills: 64% probability of detection x 60% probability of effective deterrence = 38% probability of reducing risk of entering RSZ of spinning turbines

Data for all eagles combined from the Goodnoe Hills (bald eagles rarely occur at the Manzana site) indicated similar results as for golden eagles alone, except that limited data suggested the probability of effective deterrence was higher for bald eagles than for golden eagles.

The Goodnoe Hills control-treatment experimental setup allowed for confirming that the addition of DTBird audio deterrents significantly increased the probability of effective deterrence compared to spinning turbines alone (deterrent signals muted). The difference amounted to a 1.8–2.3-fold (depending on signal type) increase in effective deterrence beyond the influence of spinning turbines for golden eagles alone, and a 2.1–2.9-fold increase for all golden and bald eagles combined, with bald eagles appearing more sensitive to the audio deterrents than golden eagles. We have no basis for comparison at the Manzana facility, but we suspect similar proportional effects would be evident there, perhaps heightened somewhat by evidence of greater overall deterrence effectiveness at that site.

Recalculating the estimates of DTBird's overall detection and deterrence effectiveness for golden eagles alone based on the added benefits estimate from the Goodnoe Hills results in the following modifications:

> Manzana: 66% probability of detection x 40% probability of added effective deterrence = 24% probability of reducing risk of entering RSZ of spinning turbines
>
> Goodnoe Hills: 64% probability of detection x 30% probability of effective deterrence = 19% probability of reducing risk of entering RSZ of spinning turbines

If we further narrow the focus to evaluating DTBird's effectiveness in detecting eagles (or UAV surrogates) and deterring eagles that were flying in core exposure locations (i.e., primary dissuasion-trigger risk zone within approximately 170 meters or less of the relevant turbines) and that we classified for behavioral analysis as at moderate to high risk of exposure to the RSZ of spinning turbines, the estimates of effectiveness across the two study sites increase markedly as follows:

Effectiveness of Spinning Turbines + Deterrents: 68% probability of detection x 80% probability of effective deterrence = 54% probability of reducing risk of entering RSZ of spinning turbines

Added Effectiveness of Deterrents: 68% probability of detection x 44% probability of effective deterrence = 30% probability of reducing risk of entering RSZ of spinning turbines

By eliminating from the equation eagles that were at low risk of approaching the RSZ of turbines and whose behavior was less likely to be influenced by either the spinning turbines or triggered audio deterrents, these heightened estimates of effectiveness are more likely to represent the true proportional benefits of the DTBird systems in reducing the risk of golden eagles entering the RSZ of focal turbines at the two study sites.

Our second approach to quantifying DTBird's overall effectiveness stems from the 2-year controlled experiment comparing eagle activity rates at DTBird installations operating in control mode with deterrents muted and in treatment mode with deterrents broadcasting normally. For golden eagles alone, the dissuasion-trigger and dwell-time models indicated similar reductions (27–29%) in indicative activity rates at turbines with the audio deterrents broadcasting compared to turbines with the audio deterrents muted. Assuming activity rates are positively correlated with the potential for collision risk, these percentage estimates of reduced activity levels in the vicinity of treatment turbines should represent roughly comparable estimates of DTBird's deterrence and collision-risk reduction benefits as those derived from our first estimation approach. Assuming this is true, the proportional estimates of collision-risk reduction from DTBird for golden eagles derived from the various estimation approaches were notably similar (19%, 27%, and 29%). Together these results suggest that, for golden eagles that fly anywhere within the calibrated maximum detection range for the species, operation of the DTBird automated detection and audio deterrence system can be expected to reduce the probability of approaching the RSZ of spinning turbines by 20–30%. Again we note, however, that further narrowing the focus to eagles (or surrogates) whose flight patterns exposed them to relatively high risk of entering the RSZ of turbines elevated the estimate of core effectiveness by at least 11%.

Properly scaled and tailored to the unique "survey" effort represented by the automated DTBird monitoring (not an easy task in this case due to highly variable turbine-specific sampling over 25 months), the dwell time data potentially could be translated to a surrogate for the pre-construction "eagle activity minutes" metrics used to project fatality rates at wind-energy facilities using the Bayesian collision risk model developed by the U.S. Fish & Wildlife Service (2013) and partners (New et al. 2015). If so, one could then theoretically compare independently projected post-construction fatality estimates tailored to the Goodnoe Hills based on dwell-time activity levels at control turbines versus treatment turbines to derive a quantitative estimate of projected fatality reduction from operation of DTBird at that facility. However, the magnitude of such a comparison (i.e., a reduced number of fatalities/year) could not be directly extrapolated to other facilities with different collision-risk infrastructure and eagle activity rates and behaviors. Instead, our perspective is that proportional/percentage estimates of effectiveness can be more easily tailored to projecting the magnitude of DTBird's beneficial effects in reducing collision risk at different facilities once initial pre-construction fatality projections tailored to the specific site are developed using the USFWS Bayesian risk model.

We designed this study to yield overarching insight about DTBird's effectiveness by sampling across an array of turbine-specific installations at two study sites, but with no expectation of producing facility-level estimates of effectiveness based on evaluating the influences of specific spatial arrays and densities of DTBird installations. As a result, the estimates of effects summarized herein should be thought of only as indicators of how individual DTBird systems can be expected to influence activity around the specific turbine on which a given system is installed. The estimated proportional effects can certainly be extrapolated across multiple turbines within a facility to develop a sense of the potential aggregate effects of installing multiple DTBird systems, but cannot be used to infer potential interactive benefits that could accrue from having multiple installations arrayed in particular configurations. Further, the comparative result we derived from the two study sites—one in a desert foothills landscape and one in temperate grassland ridgeline landscape—clearly indicated that DTBird's overall effectiveness may vary in different landscape/climatic settings with different resident and transient eagle populations and variable false-positive deterrent-triggering rates that may influence the eagle responses.

4.5.6 Performance Reliability and Cost Analysis (Objective 6)

Sixteen DTBirdV4D8 units were manufactured in 2019 and delivered to Goodnoe Hills wind farm by the end of the year. 14 units operated under the evaluation and experimental design from August 2021 to September 2023. When including the overall cost of LIQUEN´s Internal Services and R&D Department, the standard DTBirdV4D8 model sale cost (cameras model Falco and Larus software) is around $18K - $22K, and the yearly service sale cost around $2K - $3K. There are other project specific indirect costs for installation (around 4K$-6K$ per unit) and onsite maintenance (around 0.6 K$-2K $ per unit and year) (Table 27).

Table 2727. Actual Cost(s) to Install, Operate, and Maintain the DTBird system (2016-2024).

Project Cost(s)	Amount (USD)	Unitary cost for the 14 units (USD)
Actual purchase cost for 14 DTBirdV4D8 Units	$208.619,64	$14.901,40
Shipping and customs for DTBird Units to Goodnoe Hills*	$17.114,49	$1.069,66
Installation costs (travel and salary)	$10.659,23	$761,37
Year 1 service costs: 12 months of service, including travel costs to repair multiple maintenance issues August 2021 – July 2022	$42.997,43	$3.071,25
Year 2 service costs: 12 months of service, August 2022 – September 2023	$35.199,41	$2.514,24
Total	**$327.278,51**	**$23.377,04**

*16 units were delivered to the site

Ongoing technical complications reduced the team's ability to collect the intended level of data across the originally anticipated sample of turbines. PacifiCorp and Liquen experienced many technical troubles during repeated attempts to integrate installed DTBird units into the Goodnoe Hills Washington wind facility's SCADA and network.

- Four units experienced repeated camera failures.

- Persistent malfunctions on GH40 and GH41 early in the commissioning process could not be resolved and the units were permanently removed from the experiment in favor of recovering other turbines further ahead in the commissioning process and maximizing the number of turbines commissioned in time for UAV flight trials. Parts from these units were used to resolve issues in other units.
- The malfunction on GH56 could not be resolved in time for the UAV flight trials or for the unit to participate in the Year 1 experiment, but PacifiCorp and Liquen continued to troubleshoot and with an onsite visit in C22-Q3 they have resolved the issues with this unit making it available for year 2 of the experiment.
- The onsite visit also revealed that during installation, equipment for G51 was mistakenly installed in G56 and equipment for G56 was installed in G51; data collected at G51 was mistakenly assigned to G56 and vice versa. Because of this confusion, data from G51 were not available for the analysis of the Year 1 experiment or from the UAV flight trial in July 2022. We are working to determine if data collected at G56 can be retrieved and included in further analysis of the experiment data
- Additionally, a month-long power outage at the Goodnoe Hills project site resulted in communication loss between DTBird and the SCADA system. The DTBird system was not responsible for this power loss and did not necessarily affect the system's ability to function as intended in detecting target species and triggering audio deterrents, it hindered our ability to evaluate the system in real time and the cause has not yet been resolved.

While some of the challenges resulted from less-than reliable maintenance at the study site, our experience suggests that reliable use of the DTBird system requires regular service at the systems. In addition to troubleshooting malfunctions, this includes:

- Two months of refining Liquen's detection algorithm in the field before installed DTBird systems can be considered fully commissioned
- Replacing all camera lenses at DTBird systems every six months, along with quality assurance to make sure camera positions have not changed during regular maintenance.

4.6 Discussion

4.6.1 DTBird Detection Performance

The specifications promulgated by Liquen (2017) specified that DTBird systems comparable to those installed at Manzana and Goodnoe Hills should be expected to result in a yearly average TFP deterrent trigger rate of 0.2–4.0 events/turbine/day, amounting to a total duration of 0.1–2.5 minutes/turbine/day. With seven turbines evaluated across 10 months, estimates from the Manzana study fell within these ranges: averages of 1.2–1.8 TFP deterrent triggers/turbine/day among the seven turbines and an overall average rate derived from the integrated analysis presented herein of 1.3 TFP detections with deterrent triggers/turbine/day. Similarly, the TFP

deterrent emittance rate (warning and dissuasion signals combined) was estimated to average 0.6–0.9 minutes/turbine/day among the seven installations, and the overall average rate derived from the integrated analysis presented herein was 0.8 minutes/turbine/day. Results from the Manzana study and other prior studies of DTBird technology (May et al. 2012, Aschwanden et al. 2015) formed the basis for the performance targets specified for the Goodnoe Hills study: maximum of 1.6–2.8 TFP deterrent triggers/turbine/day, and no more than 36% of all relevant detection events resulting from TFPs.

The overall-average TFP deterrent-triggering event rate at Goodnoe Hills across 23 months of sampling was 3.9 TFP deterrent triggers/turbine/day, which substantially exceeded the established performance target. However, after Liquen made additional adjustments to reduce the false positive rate in January 2023, the rate for the subsequent 7 months dropped to an average of 0.8 TFP deterrent triggers/turbine/day, well below the performance target. Similarly, although TFPs resulted in more than 50% of all detections that triggered deterrents before the adjustments were made, the proportion dropped to 25% post-adjustments, again falling below the established performance target. Moreover, in both cases the post-adjustment rates at Goodnoe Hills were lower than at Manzana, suggesting improvement in the filtering algorithms.

Across the periods of record, the overall TFP-caused deterrent signal durations (warning and dissuasion signals combined) on turbine-days when deterrents were triggered averaged 0.84 minutes/turbine/day at Manzana and 0.96 minutes/turbine/day at Goodnoe Hills. Post-adjustments, the combined deterrent signal duration rate at Goodnoe Hills fell only slightly to 0.95 minutes/turbine/day, despite the significant reduction in numbers of TFPs. This suggests that fewer signals averaged longer in duration per trigger after the adjustments, which may indicate that birds exposed to fewer TFP-triggered deterrents may have subsequently lingered more around the turbines with DTBird installations (a possible manifestation of negative habituation to prior excessive TFP signaling). Regardless, all documented signal duration rates fell below Liquen's desired standard of <2.5 minutes/turbine/day.

The results focused on variation in the prevalence of TFPs caused by insects during Goodnoe Hills Year 1 suggested the potential for substantial seasonal variation at this site, with a lesser magnitude of seasonal variation also evident at Manzana. However, the comparative results for Goodnoe Hills Year 2 suggested that the additional adjustments Liquen made in 2023 substantially mitigated/dampened what would otherwise have continued to be a significant source of excessive deterrent signaling during summer/fall at Goodnoe Hills (and perhaps at Manzana had earlier adjustments not been made there).

The notable contrasts in temporal patterns of sky artifact TFPs among years at Goodnoe Hills and between Manzana and Goodnoe Hills likely reflects a combination of factors. First, the documented difference in prevalence in Goodnoe Hills Years 1 and 2, showing a similar pattern as for insect TFPs, suggested that the further adjustments to the false positive filtering algorithms Liquen made in early 2023 probably also reduced the probability of sky artifact TFPs and contributed to the much lower post-adjustments sky artifact TFP rate in Year 2 compared to the corresponding cycles in Year 1. However, examining the patterns in relation to calendar months also suggested the possibility of weather-related differences in the source of TFPs at

the two sites. Specifically, sky artifact TFPs were generally common throughout the year at Goodnoe Hills and appeared to be particularly prevalent from spring through mid-summer (in Year 1 when not limited by additional filtering), whereas sky artifact TFPs appeared to be more restricted to late winter/early spring at Manzana. This suggests that the variable climatic regimes of the two study regions also contributed to the differences in the TFP rate between the two sites. Specifically, highly dynamic, partly cloudy skies tend to be more restricted to late winter/spring in the relatively xeric environment of the Mojave Desert where the Manzana site lies, whereas variable storminess and cloudy weather are often consistently more prevalent both during snowy winters and extending later in spring and into early summer in the Columbia Gorge region of Washington where the Goodnoe Hills site lies. Sky-artifact TFPs appear to arise more frequently when cloud cover is more prevalent and variable, dynamically producing more high-contrast elements that the DTBird system erroneously interprets as target movement.

Data from the Manzana and Goodnoe Hills study sites were also similar in showing some common species and seasonal patterns in the prevalence of detections reflecting the activities of NTAFPs. Common ravens were the most common source of NTAFPs at both sites, with generally higher activity during spring and fall migration, lowest activity during mid-summer in California, and moderate activity during winter in both areas.

Excessive false-positive detections hinder effective use of the DAP system for tracking activity and identifying exposure risk for focal species. This required investigators to sift through thousands of false positive records that did not trigger deterrents when the study motivation calls for screening such records (H. T. Harvey & Associates 2018). While burdensome for the purposes of this study, we did not find that excessive FPs led to negative habituation during the study period. Instead, we found evidence that excessive FPs may have led to positive habituation and potentially decreased risk to eagles. However, this does not address other potential burdens associated with FP detections excessively triggering deterrents, including potentially negative consequences of sound pollution to non-target wildlife and personnel. More generally, the results of this study clearly illustrate that limited AI discernment capabilities combined with audio deterrents may result in variable system effectiveness.

The probability of detection/false negatives models, resulting from the UAV flight trials, indicated similar patterns at the two study sites, with a nominally higher detection probability at Manzana (66%) than at Goodnoe Hills (64%). These estimates exceed the performance standard of 63% established as a basis for evaluating DTBird performance at Goodnoe Hills, though clearly nothing done to potentially improve the detection systems between the Manzana pilot study and the subsequent Goodnoe Hills study led to better performance at Goodnoe Hills. Instead, this outcome suggests consistent performance of the primary detection functions of the DTBird systems at both sites.

The probability of detection modeling analysis also provided useful perspective concerning factors that influence the overall probability of DTBird detecting an eagle-like UAV if it flies anywhere through the detection envelope projected based on calibration for golden eagles. The limitation of this analysis is that for flights that are not detected (false negatives) there are no reference points to use for precisely characterizing the flight, location, and environmental

characteristics at the time of a specific DTBird event to use as covariates. Consequently, we focused attention on discerning the influences of only a select few metrics derived by using GIS tools to calculate selected minimum and averaging position metrics across all points along a given sample flight. Nevertheless, this relatively simple approach illustrated variability in the probability of detection through the day, likely related to the relative influence of solar position and intensity.

More importantly, the results emphasized that the probability of detection was highest when the target flew at moderate distances from the turbine (generally high with average flight distances of 80–160 m) through the midsection of the camera viewshed (generally high with viewing angles from camera up to UAV of 25–40°). Conversely, the probability of detection averaged lower when the target flew either closer to or farther away from the camera or primarily within the lower or upper margins of the camera viewshed. These results are perhaps not surprising in suggesting that detection tends to be lower around the margins of the camera viewsheds and higher when a bird is flying at moderate distances from and in the center of a camera viewshed. The latter conditions are exactly when birds approaching a spinning turbine tend to be at greatest risk of entering the RSZ of spinning turbines. However, especially hunting or displaying raptors such as golden eagles often make very dynamic movements that can either rapidly drop them down from up high or pop them up from down low and quickly bring them into the RSZ danger zone at relatively close range. For this reason, poorer detection low and close or high and close to the turbine can result in problematic interactions with little time for the deterrents to trigger and discourage continued closer passage before entering the collision risk zone.

Characterizing the response-distance data for the three event types (detection, warning signal, and dissuasion signal) revealed some unexpected results. The average response distance for detection events (190 m) was longer than for dissuasion signals (176 m), as expected, but was considerably shorter than the 240-m theoretical maximum, calibrated detection distance. This result primarily reflected that initial detections often occurred when the UAV flew in low and first entered the detection envelope from the underside of the overall, inverted-cone-shaped envelope at relatively close distances to the turbine. Conversely, longer-than-expected response distances were comparatively uncommon.

A similar factor also contributed to the outcome for warning signals, where some initial triggers were expected to occur at distances of 100–170 (Figure 4); however, with the realm over which such warning signals could occur limited to less than one third of the perimeter area over which shorter detection distances could arise (Figure 4), the matching average detection and warning signal response distances were not expected. Reasons for this result are uncertain, but the outcome may reflect that, despite mostly common triggering calibration, longer than expected warning-signal response distances were proportionately less common than longer-than-expected detection response distances. This could be considered a desirable outcome, in that it means relevant targets were sometimes detected at greater than expected distances—increasing time for effective deterrent response if needed—but unnecessary warning signals targeting extra-distant birds were constrained.

The average response distance for triggering a dissuasion signal (176 m) nearly matched the calibrated core-envelope trigger distance for that event type (170 m), whereas the expectation was for a lower average reflecting a mix of expected response distances of approximately 170 m across the core-envelope surveillance area and 100-m in the outer, lower band of surveillance areas (see Figure 4). Instead, the observed outcome suggested that dissuasion signals were triggered more often than expected at distances exceeding the calibrated trigger distances. This result could be considered a beneficially conservative outcome in providing more time for an approaching bird to respond to a dissuasion signal, as long as it does not result in unnecessarily excessive triggering of the signals, with possible adverse consequences for non-target wildlife, facility staff, or facility neighbors (H. T. Harvey & Associates 2018).

The multi-site results illustrated notable random variation among turbines at the two study sites, and indicated that, given modeling of other random and fixed effects, the overall DTBird response distances tended to average marginally shorter at the Manzana study site compared to the Goodnoe Hills site. Reasons for this difference are uncertain, but it suggests that the overall targeting accuracy of the DTBird systems can vary slightly across different landscape settings, perhaps reflecting inherent differences in the overall visual clarity and complexity of different regional skies and landscape backdrops. DTBird does not reliably detect objects against a landscape, as opposed to sky, backdrop, and topographic complexity sometimes intrudes within the camera viewsheds to limit detectability. In this case, the proximate and elevated backdrop of the Tehachapi Mountains may have complicated detectability at the Manzana site more than the comparatively wide-open skies at the Goodnoe Hills site.

The multi-site results continued to support the notion that modeled variation in average response distances among the five UAV models we deployed in this study likely mimicked the kind of random variation that could be expected given eagles of different sizes and coloration patterns, such as those pertaining to differences among the sexes and age classes of golden eagles. As the initial Manzana site-specific analysis suggested (H. T. Harvey & Associates 2018), the demonstration that response distances tended to be relatively short for the AUV Custom aircraft is logical given its skinny tubular hind body and overall modest stature, with the relatively long-winged but slender Ranger aircraft also showing some of that tendency. In contrast, a tendency toward longer response distances was associated with the overall more eagle-like and robust-bodied AES Custom and Clouds models.

The multi-site results pertaining to the influence of cloud cover / sky backdrop on DTBird response distances suggested some similar patterns as the preceding site-specific analyses, but also some refinements. Specifically, all else equal, the updated analysis indicated that response distances generally increased as cloud cover increased and averaged significantly longer once the cloud cover extended throughout the viewshed under relatively uniform overcast skies. This outcome is logical in suggesting that the DTBird systems more readily detected the relatively dark eagle-like UAVs against relatively uniform high-contrast white or gray backgrounds than against less contrasting blue skies and or highly dynamic partly cloudy skies. These tendencies also mimic the challenges faced by observers scanning the skies for migrating raptors, where the presence of uniform cloud cover greatly increases the detectability of migrants passing overhead underneath the clouds (Bildstein et al. 2007).

The multi-site model uniquely indicated a significant positive association between response distances and UAV ground speed, which suggested that targeting performance improved significantly when a UAV was traveling relatively quickly from the perspective of the camera. This result may reflect that the DTBird detection algorithm focuses on targeting objects that both fill enough image pixels to warrant targeting from an estimated distance perspective, and that it perceives as moving in a manner that could be a flying bird. Our modeling results suggest that, across the UAV flight speeds documented in this study, slow-moving targets were generally harder for the DTBird system to detect than rapidly moving targets.

We included in our modeling effort consideration of a suite of variables as potential indicators of variation in the exposure of UAV profiles to the cameras, where greater profile exposure is expected to increase the accuracy of DTBird targeting based on calibrated settings. Our hypothesis was that the more a UAV climbs or descends, pitches up or down in the wind, rolls from side to side in the wind or while banking, or is generally bounced around by and quarters into the wind, the more the UAV profile should be exposed to the cameras and lead to more accurate targeting. Similar to the preceding site-specific modeling results, the final multi-site model continued to emphasize the importance of such variables in predicting DTBird response distances—specifically indicating a positive association with wind speed and the interactive influence of roll and pitch angels. The previous site-specific models also suggested that UAV *Climb Rate* was a relevant predictor, but that variable did not pan out as a significant predictor in the multi-site model, perhaps due to the combined data reflecting a stronger association with pitch and roll angles, with the former variable theoretically capturing a similar effect as variable climb rates (both descending and descending trajectories). The final model indicated relationships for wind speed and roll and pitch angles that were similar to the patterns reflected in the previous site-specific models, suggesting that response distances increased at higher wind speeds (UAV bouncing around more) and/or when the UAV was rolling side to side more, but only if the aircraft was not simultaneously pitching up or down to a substantial degree, because that combination would have caused the aircraft to stall and fall from the sky.

The initial Manzana site-specific model reflected a significant second-order relationship between response distances and the intensity of solar irradiation impinging on the UAV in the direction of the cameras. However, that relationship did not pan out again in the Goodnoe Hills site-specific model once we applied a more robust approach to developing that model. Similarly, none of the solar variables we considered were incorporated in the final multi-site model. There is no question that flying objects seen in the DTBird videos and targeted by the system routinely disappear from view when passing through major sunspots, and that high intensity solar insolation often increases the glare factor around such sunspots. In this case, however, we suspect that the combined-site dataset more effectively captured this effect in the refined relationship with sky backdrop/cloud cover. Specifically, situations where substantial sunspots obscured detectability were particularly prevalent under fair and partly cloudy skies, and greatly diminished when cloud cover was more complete, especially once overcast skies prevailed. Hence, the relative prevalence of sunspots may have been a primary driver behind the apparent positive relationship between response distances and cloud cover illustrated in the

multi-site model, to the exclusion of solar intensity or positioning proving to be of additional predictive value.

This investigation highlighted several flight metrics and environmental covariates that significantly influenced DTBird's detection and deterrent-triggering performance at the two wind-facility study sites. Here it is important to acknowledge that using eagle-like UAVs as surrogates for real eagles constrained the insights generated from the study. We think the fixed-wing UAVs we used in the study did a good job of mimicking the non-flapping soaring and other flights of eagles, but were limited by not having wings that flap and tuck in the manner used by eagles to accomplish various maneuvers. The UAVs were also not capable of undertaking steep dive-and-roll or "roller-coaster" type display maneuvers that Golden Eagles sometimes make in pursuing prey or as part of their territorial behavior (Katzner et al. 2020). The degree to which more-dynamic wing action and flight maneuvers could alter the apparent targeting performance of the DTBird systems is uncertain. Wing flapping undoubtedly exposes more of a bird's profile to the cameras, at least intermittently; however, wing tucking does the opposite. In other words, these two components of real-bird flight dynamics may be offsetting factors that translate to average response distances similar to those reflected in the strictly fixed-wing UAV data we collected. If efforts to use UAVs as bird mimics are considered for similar future studies, some of the new robotic birds available today that actually fly with flapping wings should be considered, as long as the flapping rate of the robotic bird effectively mimics that of target birds of interest. In particular, a robotic bird with quick wingbeats and that flaps all the time to stay aloft would not be a good mimic for eagles, because eagles often spend most of their time in non-flapping soaring and sailing flight, rather than using powered flight (e.g., see Katzner et al. 2020).

Throughout these UAV flight trials, our effort was unexpectedly constrained to a high degree by incompatible weather and wind conditions. High winds and excess moisture in the air not only limited when we could fly, but also ultimately led to fatal crashes that took out four of the five aircraft we used, because we were compelled to fly in conditions that pushed the limits of tolerance for the light-weight, foam-bodied aircraft. On the positive front, having to replace several aircraft resulted in our flying a greater diversity of models than initially anticipated, which effectively mimicked some of the variability in DTBird performance that would likely occur given eagles of various sizes and color patterns. On the negative front, these unexpected complications significantly reduced the diversity of flight conditions during which we were able to conduct sampling flights, and substantially constrained the overall dataset compared to our original study-design projections. Nevertheless, we think the dataset we did amass provided valuable insight into how salient flight characteristics and environmental covariates influenced DTBird's performance in detecting eagles (or surrogates) and triggering deterrence signals compared to calibrated system settings.

Lastly, we acknowledge that the differences rated as statistically significant effects given our data sometimes amounted to effects magnitudes that may not have especially noteworthy biological or operational significance (e.g., 10–20 m differences in detection range for birds that may easily move farther than that in less than a second). However, our study was not designed to specifically quantify the relative effectiveness of different calibrated detection and deterrent

triggering distance thresholds nor the spatiotemporal aspects of what an eagle requires as deterrent warning to avoid calamity under different flight conditions. Therefore, we have no firm basis for presuming what may be biologically/operationally significant in this context.

4.6.2 Behavioral Differences at Treatment vs Control Turbines

The *in situ* two-year experiment at Goodnoe Hills, Washington failed to reveal a significant overall treatment effect on the probability of a target bird triggering a dissuasion signal (Hypothesis A), but did reveal an effect of treatment on dwell time (Hypothesis B).

One possible reason the former relationship was not apparent concerns the efficacy of warning signals as a potential means to reduce the probability of an eagle triggering a subsequent dissuasion signal. Although eagles triggered warning and dissuasion signals with similar frequencies overall, a large majority of the triggered dissuasion signals were not preceded by a prior warning signal. In other words, the idea that broadcasted warning signals could be expected to reduce the probability of triggering a subsequent dissuasion signal actually did not apply very often. Two potential explanations for this pattern are: 1) within the primary detection envelope where sequential warning and dissuasion signaling is expected when relevant, the DTBird detection systems frequently did not detect eagles until they had already reached the closer dissuasion-triggering envelope; and 2) eagles often flew in relatively low and entered the detection envelope relatively close to the turbine where dissuasion signals were immediately triggered without a prior warning signal.

The significant effects of *Treatment Group* in the dwell-time models translated to predictions of golden eagles and all eagles combined averaging 24–27% less time dwelling in the vicinity of DTBird systems operating with their deterrents broadcasting normally compared to systems with muted deterrents. The golden eagle dissuasion-trigger model indicated a similar—albeit statistically nonsignificant—29% decrease in the probability of dissuasion triggers at treatment turbines. Quantifying estimated reductions in the probability of dissuasion triggers at treatment turbines based on the all-eagles model was complicated by the presence of interactions with both categorical (*Cloud Cover*) and continuous (*FPs per Day*) covariates. Under most sky conditions from partly cloudy to overcast, eagles tended to trigger approximately 9–30% fewer dissuasion signals at turbines with DTBird deterrents broadcasting normally (i.e., consistent with research Hypothesis A), whereas a much stronger, opposite pattern was shown when fair skies prevailed. Reasons for this unexpected anomaly are uncertain, but one possibility is that visibility typically tends to be clearer overall during fair weather. Better visibility might have allowed the eagles to more easily perceive the spinning turbines, take heed of the broadcasting deterrents, but also remain more comfortable flying and foraging closer to the turbines with less concern for the potential collision risk. In contrast, the indicated interactive relationship between *Treatment Group* and *FPs per Day* indicated further clear support for Hypothesis A in demonstrating that the positive effect of broadcasted deterrents at treatment turbines deterring eagles from triggering dissuasion signals was accentuated by higher FP deterrent-triggering activity, whereas no such effect was evident at control turbines. The difference in the probability of dissuasion triggers at control versus treatment turbines was nominal when the FP deterrent

triggering rate was low, but was approximately a 60% lower at treatment turbines when the FP deterrent triggering rate was elevated.

The model focused on presumed golden eagles triggering dissuasion signals indicated a novel relationship with monitoring *Year* as a predictor, suggesting that the probability of golden eagles triggering dissuasion signals declined overall by approximately 46% across the facility during Year 2 of the study. Neither *Year* nor *28d Cycle* emerged as a significant predictor in the all-eagles dissuasion-trigger model; however, *28d Cycle* emerged as an important predictor in the dwell-time models for both golden eagles alone and all eagles combined. Similar to the result for golden eagles and dissuasion triggers, the indicated relationship for *28d Cycle* was an overall declining trend across the 2-year study in the dwell time of golden eagles alone and all eagles combined. Given that these trends did not emerge differentially around DTBird equipped turbines operating in treatment versus control mode, the overall pattern may provide evidence of positive habituation through time among resident and seasonally resident eagles. As such eagles became increasingly exposed to deterrents being broadcasted regularly around the perimeter of the facility, they might have grown increasingly wary of dwelling for extended periods in the vicinity.

Here it is important to note that this potential habituation pattern could have been accentuated by two factors: 1) an unusually high overall FP triggering rate through the first 19 months of the study, until Liquen was authorized to undertake further fine-tuning of the filtering algorithms to reduce the FP rate; and 2) due to an extended failure of communications between the DTBird and turbine SCADA systems following a forced 24-day site-wide power outage, all DTBird systems operated in default mode after May 2023, whereby the deterrents were being triggered whether or not the focal turbine was spinning. The first factor substantially reduced the overall FP deterrent triggering rate after January 2023; however, the second factor may have largely offset that effect by increasing the overall prevalence of superfluous deterrent triggering in after May 2023. This combination likely maintained an elevated rate of deterrent triggering throughout most of the 2-year study, which could have accelerated the pace of any positive habituation effects. What is equally important to note here, though, is that the results do not point to possible negative habituation, which would involve eagles learning to ignore the deterrents and remain at risk.

All of the models we developed reflected a pronounced diel pattern of variation in the documented eagle responses that operated independently of the applied deterrent treatment regime. Most of the modeled results captured the relationship as increasing strongly—whether the probability of dissuasion triggers or average dwell time—from dawn until reaching a mid-afternoon peak, followed by a lesser, gradual decline until dusk. We think this predominant pattern probably reflects the common general activity levels of eagles and other raptors during a typical day, with the flight activity of especially large soaring raptors typically dependent on thermal and wind activity increasing as the day warms up to provide energy-saving lift for active foraging, patrolling, and other flight-dependent activities.

Finally, *Wind Speed* emerged as significant covariate influencing the probability of golden eagles triggering dissuasion signals, independently of the implemented control-treatment design. The

indicated effect of higher wind speeds generally reducing the probability of dissuasion triggers suggests that the faster the turbines are spinning the more they themselves act as a deterrent to visually acute eagles, who then remain farther away from the perceived danger independent of the influence of DTBird deterrent signaling.

4.6.3 Behavior Responses Across Both Sites

The results of the multi-site, integrative analysis of large-raptor behavioral responses to broadcasted DTBird audio deterrents illustrated noteworthy differences in the apparent responsiveness of golden eagles, turkey vultures, and buteos at the two wind facilities located in different landscape settings. When exposed to broadcasted deterrents, on average, the birds at the Manzana facility in a California foothills/desert landscape appeared to respond more effectively than their counterparts at the Goodnoe Hills facility occupying a ridgetop/grassland landscape bordering the Columbia River in Washington. Reasons for this difference are uncertain but could reflect the influence of differences in the relative proportions of different species and residents versus transients frequenting the two sites, with variable sensitivities and habituation tendencies. Alternatively, variable wind and climate regimes may have differentially influenced the response behaviors of birds at the two sites by influencing birds' abilities to hear and respond to the deterrents. Wind speeds recorded as part of the records analyzed for this analysis averaged and gusted slightly higher at the Goodnoe Hills (average 6.3 ± SD of 3.41 m/s, maximum 21.1 m/s) than at the Manzana site (average 5.7 ± 2.79 m/s, maximum 17.0 m/s); however, the modeling results suggested that higher wind speeds tended to increase rather than decrease the probability of effective deterrence. Note, however, that eagles tended to be increasingly more responsive to the deterrents than vultures and buteos as wind speeds increased, and there was some suggestion that, for golden eagles, the probability of effective deterrence tended to be higher at the Goodnoe Hills than at the Manzana site at moderate and higher wind speeds. These tendencies may have helped to ameliorate the evident site-specific difference in deterrence effectiveness during periods of high wind speeds and power production at the Goodnoe Hills. Regardless, the documented site differences clearly suggest that effectiveness of the DTBird deterrence system may vary significantly depending on the local landscape characteristics and species assemblages.

Both the multi-species and golden eagle models also reflected at least marginally significant relationships between the probability of deterrence and wind speed. Increasing wind speeds generally resulted in a higher probability of effective deterrence for larger eagles and vultures, but not for smaller buteos. We included wind speed as a potential predictor in the LGLMs thinking that higher wind speeds could reduce the probability of effective deterrence by either limiting a bird's ability to hear the deterrents and/or hindering its ability to maneuver effectively in response to the deterrents. The modeling results suggested our hypothesis was incorrect, however, at least for the larger eagles and vultures. One possibility is that faster-spinning turbine blades themselves act as a greater deterrent to approaching larger birds and more effectively amplify the effect of the audio deterrents. It is also possible that higher wind speeds actually facilitate greater maneuverability and responsiveness in many cases for large soaring raptors, which often strongly rely on the energy savings provided by wind-driven (or thermal) lift. In contrast, smaller buteos are generally more maneuverable and more easily constrained by

strong winds, such that increasing wind speeds may be a detriment rather than a benefit for them in influencing their ability to respond effectively to the deterrents.

Evidence that the probability of effective deterrence tended to be highest for birds we classified as at moderate risk of exposure to turbine collisions, rather than for those we classified as high risk of exposure, also may relate to birds having enough time and room to maneuver effectively in response to the deterrents. We expected responsiveness to be lower for birds at low risk of exposure, because such birds have little need to divert their flights to avoid risk. In contrast, birds at high risk of exposure may appear less responsive simply because they have less time and room to respond effectively if not deterred before entering a high-risk zone.

Accurately characterizing the behavioral responses of raptors to the DTBird audio deterrents was greatly confounded by two primary factors: 1) low-resolution video recordings frequently obscured the details of bird behaviors, such as changes in flapping rates, distinct “flinches” and head movements, and subtle flight path alterations; and 2) seeking insight about the degree of response based on evaluating two-dimensional renderings of three-dimensional movement scenarios, especially pertaining to measuring flight diversion angles as a relevant criteria. With this perspective in mind, if eagles and other raptors tended to respond to the deterrents less dramatically, but nonetheless effectively, at the Goodnoe Hills, then the limitations outlined above could have more easily reduced our ability to effectively discern subtler effective responses at the Goodnoe Hills. For this reason, comparing the proportions of only confirmed effective responses at the two sites may be misleading, as opposed to focusing on the combination of effective and potentially effective responses as a better comparative indicator of relative success.

The Goodnoe Hills results clearly did not meet the performance metric established based only on confirmed effective responses from the Manzana study. Further, combining CE and PE responses reduces but does not eliminate the indication of greater deterrence effectiveness at the Manzana facility, but it does result in effectiveness metrics for both sites and all species groups that exceed the ≥50% effectiveness threshold established as performance metric for this DOE-sponsored research project (Table 14). Taking this approach may overestimate DTBird’s effectiveness to some degree. We expect, however, that there is a higher likelihood of underestimating the system’s effectiveness by limiting the results to confirmed effective responses, because of our limited ability to confidently discern and classify relatively subtle but nonetheless effective behavioral responses.

The control-treatment setup for the Goodnoe Hills study provided further insight about the degree to which responses to spinning turbines and broadcasting audio deterrents contributed to the effectiveness statistics presented herein. Based on the comparative control-treatment results and for all analyzed groups and species, broadcasted deterrents consistently resulted in at least a doubling of the proportion of cases where an effective or potentially effective response was evident. Further, results for all four analyzed species groups consistently indicated that confirmed effective responses were more common when the deterrent signals were broadcasting, and that birds exhibiting no apparent response at the time a deterrent was triggered were always significantly more common when the deterrents were triggered only

virtually. However, we had no ability to conduct a similar control-treatment evaluation at the Manzana site to provide comparatively robust insight to determine if a similar proportional effect of spinning turbines and broadcasted deterrents would apply at the two sites.

In summary, the results of this investigation pointed to noteworthy differences in the apparent effectiveness of the DTBird deterrence system in different landscape settings, for undetermined reasons but with species and wind-regime differences potentially important. The results from the Goodnoe Hills site in Washington suggested a lower level of deterrence than the results from the Manzana site in California, which fell well below the ≥50% effective deterrence performance standard, when including confirmed effective responses alone. However, when considering both confirmed and potentially effective behavioral responses, the probability of effective deterrence given broadcasted deterrents exceeded the established performance standard for golden eagles at both the Manzana (79%) and Goodnoe Hills (61%) sites, with similar results obtained for the multi-species group and vultures and buteos as independent comparative groups.

4.6.4 Eagle Collision Risk Reduction

The overarching goal of this research has been to evaluate the effectiveness of DTBird in detecting and discouraging especially golden eagles (*Aquila chrysaetos*), but also bald eagles (*Haliaeetus leucocephalus*) and other large soaring raptors from approaching the rotor swept zone (RSZ) of operating wind turbines. We initially intended to translate our results to applying the Bayesian collision risk model (CRM) recommended by the U.S. Fish and Wildlife Service (2013; and see New et al. 2015), using eagle flight times recorded by DTBird at control and treatment turbines as a proxy for eagle activity. However, we found comparisons of proportional responses to be most germane, because any estimates we could generate portraying absolute reductions in the number of eagles killed per year would be site specific, whereas proportional estimates have the potential to be applied across sites based on site-specific fatality projections. Therefore, we sought to estimate DTBird's overall effectiveness in reducing the risk of eagles entering the RSZ of spinning turbines, based on multiple complementary approaches.

The first approach involved combining probability of detection estimates derived from the UAV flight trials with probability of effective deterrence estimates derived from the behavioral analyses. The multiplicative combination of these estimates yielded an estimated 52% reduction in the probability of confirmed golden eagles entering the RSZ of spinning turbines with broadcasted deterrents at the Manzana facility, and a 38% reduction at the Goodnoe Hills facility. Data for all eagles combined from Goodnoe Hills (rare occurrences of bald eagles at Manzana) revealed similar results for golden eagles alone, except limited data suggested effective deterrence was higher for bald eagles than for golden eagles.

The Goodnoe Hills control-treatment experimental setup confirmed the addition of DTBird audio deterrents increased the likelihood of effective deterrence compared to just spinning turbines alone with deterrent signals muted. Recalculating the estimates of detection and deterrence effectiveness for golden eagles alone based on the Goodnoe Hills control-treatment results yielded a 24% probability of DTBird audio deterrents reducing risk of entering the RSZ of spinning turbines at Manzana and 19% for Goodnoe Hills. Narrowing the focus further to

estimating DTBird's effectiveness when an eagle-surrogate UAV was flying in core exposure locations and *in situ* eagles were classified for behavioral analysis as at moderate to high Preexposure Risk revealed that spinning turbines plus deterrents resulted in a 68% probability of reduced risk, with the added effectiveness of deterrents alone reducing estimated risk by 37%.

The second approach used to estimate risk reduction from DTBird was based on the Goodnoe Hills 2-year control-treatment experiment involving randomized daily rotations of muted and broadcasted deterrents. For golden eagles alone, the dissuasion-trigger (dependent variable = probability of triggering a dissuasion signal) and dwell-time (dependent variable = eagle dwell time as reflected in extent of video recording) models yielded similar estimated reductions (27–29%) in the two dependent variables at DTBird-equipped turbines when the audio deterrents were broadcasted compared to when the deterrents were muted. Combining insight from both approaches suggested that, for golden eagles that fly within the calibrated maximum detection range for the species, operation of DTBird can be expected to reduce the overall likelihood of approaching the RSZ by 20–30%, with that estimate potentially further elevated to near 40% for birds at moderate to high Preexposure Risk of entering the RSZ.

The dwell time data could potentially be used as a surrogate for the pre-construction "eagle activity minutes" metric used to project fatality rates at wind-energy facilities using the Bayesian collision risk model developed by the U.S Fish and Wildlife Service. We could have independently compared projected post-construction fatality estimates tailored to the Goodnoe Hills based on dwell time at control turbines versus treatment turbines to create an estimate of fatality reduction. However, a comparison (# of fatalities/per year) of that scale could not be extrapolated to other facilities with different collision risk infrastructure and eagle activity rates and behaviors. Therefore, we determined a better approach was to present percentage estimates of DTBird's beneficial effects in reducing post-construction collision risk, which could potentially be tailored to match initial pre-construction facility projections tailored to specific sites using the USFWS Bayesian risk model. The results from the two study sites—one in a desert foothills landscape and one in temperate grassland ridgeline landscape—clearly indicated that DTBird's overall effectiveness may vary in different landscape/climatic settings with different resident and transient eagle populations, and variable false-positive deterrent-triggering rates that may influence the eagle responses.

4.7 Conclusion

Despite falling well short of our intended 2-year sampling design due to factors beyond our control, the results of our careful analyses yielded noteworthy insight about the factors affecting the ability of the DTBird deterrent system to reduce the activity of eagles around turbines where the deterrents were broadcasting normally. Particularly notable were indications of possible long-term positive habituation reducing the dwell time of eagles around the DTBird turbines independent of the control-treatment experimental design, likely reflecting the overarching influence of an atypically elevated overall deterrent triggering rate across the installed DTBird systems. We suspect that, had frequent operational failures not caused major unexpected imbalances in our intended sampling design and had the overall deterrent triggering

not been artificially elevated by various factors, our ability to demonstrate conclusive patterns of interest concerning the proximate effectiveness of DTBird would have been even greater.

Efficiently focusing a deterrent system such as DTBird on specific species of conservation interest is often the primary objective for facility managers. In this context, avoiding unnecessary detections and deterrent signaling caused by non-focal bird species will often be important to minimize the potential risk of negative habituation.

Natural seasonal cycles in the distribution and abundance of insects contributing to TFPs and birds contributing to NTAFPs are expected, but may also occur relative to sky artifacts as solar and cloud cover variations greatly influence that source of TFPs. If predictable enough through time, it may be possible to improve the DTBird false-positive filtering algorithms to be more sensitive to these factors and thereby efficiently reduce the overall false positive rate.

Collectively, our results suggest the following should be considered in future DTBird applications:

- DTBird systems should not be considered fully commissioned and maximally effective until at least 2 months after Liquen declares the systems “commissioned” and they complete fine-tuning to minimize false positives caused by spinning blades and other factors.
- Liquen should prioritize additional improvements of the DTBird filtering algorithms to further reduce the potential for especially blade-related, insect, and sky-artifact TFPs, which result in substantial clutter within the DAP and unnecessarily trigger an abundance of potentially deleterious deterrent signals.
- Liquen should develop and implement AI systems better able to distinguish target species. NTAFPs represent a complicated management issue, in that protecting all native bird species from unnecessary human-caused mortality is a worthy objective, but excessive deterrent triggering by nontarget birds could also have negative consequences.
- Regular replacement of camera lens cover to avoid solar degradation that can radically effect the clarity of the recorded videos
- Potentially use a higher resolution camera system and sophisticated AI/ML algorithms to obviate the need to manually screen the recorded videos to identify and enumerate detected targets and evaluated their behavior. This is especially necessary when evaluating the technology, but could have additional benefits as well.

4.8 Literature Cited

ArduPilot Dev Team. 2021. Mission Planner. Version 1.3.77. https://firmware.ardupilot.org/Tools/MissionPlanner. Accessed July 2022.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67:1-48. DOI 10.18637/jss.v067.i01.

Bildstein, K. L., J. P. Smith, and R. Yosef. 2007. Migration counts and monitoring. Pages 101–116 *in* D. M. Bird and K. L. Bildstein (Editors), Raptor Research and Management Techniques. Hancock House Publishers, Blaine, WA, USA, and Surrey, British Columbia, Canada.

Bolker, B. M. 2023. Package lme4: linear mixed-effects models using 'eigen' and S4. Version 1.1-35.1 https://www.rdocumentation.org/packages/lme4. Accessed December 2023.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24:127–135.

Brooks, M., B. Bolker, K. Kristensen, M. Maechler, A. Magnusson, H. Skaug, A. Nielsen, C. Berg, and K. van Bentham. 2023. Package'glmmTMB'. Version 1.1.8. https://https://www.rdocumentation.org/packages/.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. 2017a. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal 9:378–400.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. 2017b. Modeling zero-inflated count data with glmmTMB. BioRxiv:132753. https://www.biorxiv.org/content/biorxiv/early/2017/05/01/132753.full.pdf. glmmTMB/versions/1.1.8. Accessed December 2023.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second edition. Springer, New York, NY.

Burnham, K. P., and D. R. Anderson. 2010. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second Edition. Springer, New York, NY, USA.

De Lucas, M., G. F. Janss, D. P. Whitfield, and M. Ferrer. 2008. Collision fatality of Raptors in wind farms does not depend on Raptor abundance. *Journal of Applied Ecology*, *45*(6), 1695–1703. https://doi.org/10.1111/j.1365-2664.2008.01549.x.

Erickson, W. P., G. D. Johnson, D. P. Young, Jr.. 2005. A summary and comparison of bird mortality from anthropogenic causes with an emphasis on collisions. Pages 1029-1042 *in C. J.* Ralph, C. and T. D. Rich (Editors). Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference. 2002 March 20-24; Asilomar, California, Volume 2 Gen. Tech. Rep. PSW-GTR-191. Albany, CA: U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Research Station.

Fernie, A. 2012. TLogDataExtractor. Version 1.1. https://onedrive.live.com/?cid=2B4DC8F6BD223F57&id=2B4DC8F6BD223F57%212440. Accessed August 2022.

Fisher, N. I. 1995. Statistical Analysis of Circular Data. Cambridge University Press, Cambridge, UK.

Friendly, M., and D. Meyer. 2016. Discrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data. CRC Press, New York, New York, USA.

Hair, J.F., R. L. Tatham, and R. E. Anderson. 1998. Multivariate Data Analysis. Fifth edition. Prentice Hall, Englewood Cliffs, NJ, USA.

Hartig, F. 2019. Package DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. Version 0.2.6. https://cran.r-project.org/web/packages/DHARMa/index.html. Accessed August 2022.

Hartig, F. 2021. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.4. https://CRAN.R-project.org/package=DHARMa.

Hanagasioglu, M., J. Aschwanden, R. Bontadina, and M. de la Puente Nilsson. 2015. Investigation of the Effectiveness of Bat and Bird Detection of the DTBat and DTBird Systems at Calandawind Turbine. Final Report. Prepared for the Swiss Federal Office of Energy Research Program, Wind Energy and Federal Office for the Environment Species, Ecosystems, Landscapes Division. Bern, Switzerland.

Hosmer, D. W., and S. Lemeshow. 1989. Applied Logistic Regression. John Wiley & Sons, New York, NY, USA.

H. T. Harvey & Associates. 2018. Evaluating a Commercial-Ready Technology for Raptor Detection and Deterrence at a Wind Energy Facility in California. Research report. Los Gatos, California. Prepared for the American Wind Wildlife Institute, Washington, D.C.

Hunt, G. 2002. Golden Eagles in a Perilous Landscape: Predicting the Effects of Mitigation for Wind Turbine Blade-Strike Mortality. University of California-Santa Cruz. Prepared for the California Energy Commission, Public Interest Energy Research, Sacramento, CA, USA.

Jeffreys, H., 1939. Theory of Probability. Oxford University Press, Oxford, UK.

Katzner, T. E., M. N. Kochert, K. Steenhof, C. L. McIntyre, E. H. Craig, and T. A. Miller. 2020. Golden Eagle (*Aquila chrysaetos*), version 2.0. *In* P. G. Rodewald and B. K. Keeney (Editors), Birds of the World. Cornell Lab of Ornithology, Ithaca, NY, USA. https://doi.org/10.2173/bow.goleag.02.

Length, R. 2023. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.6. https://CRAN.R-project.org/package=emmeans.

Lüdecke, 2022. Package ggeffects: create tidy data frames of marginal effects for 'ggplot' from model outputs. Version 1.1.3. https://https://cran.r-project.org/web/packages/ggeffects/index.html.

Lüdecke, D. 2023. sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.14. https://CRAN.R-project.org/package=sjPlot.

Magnussen, A., H. J. Skaug, A. Nielsen, C. W. Berg, K. Kristensen, M. Maechler, K. J. van Bentham, N. Sadat, B. M. Bolker, and M. E. Brooks. 2022. Package glmmTMB: generalized linear mixed models using template model builder. https://CRAN.R-project.org/package=glmmTMB.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. Pages 105–142 *in* P. Zarembka, editor, Frontiers in Econometrics. Academic Press, New York, USA and London, UK.

May, R., Ø. Hamre, R. Vang, and T. Nygård. 2012. Evaluation of the DTBird Video-System at the Smøla Wind-Power Plant: Detection Capabilities for Capturing Near-Turbine Avian Behaviour. NINA Report 910. Norwegian Institute for Nature Rese arch, Trondheim, Norway.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4:133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.

New, L., E. Bjerre, B. Millsap, M. C. Otto, and M. C. Runge. 2015. A collision risk model to predict avian fatalities at wind facilities: An example using golden eagles, Aquila Chrysaetos. PLOS ONE, 10(7). https://doi.org/10.1371/journal.pone.0130978.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Smallwood, K. S. 2013. Comparing bird and bat fatality-rate estimates among North American wind-energy projects. *Wildlife Society Bulletin*, *37*(1), 19–33. https://doi.org/10.1002/wsb.260.

Shah, M. 2023. Reply to inquiry: Could someone explain me about Nagelkerke R Square in logit regression analysis? Pakistan Institute of Development Economics. ResearchGate post, February 20, 2023. https://www.researchgate.net/post/Could_someone_explain_me_about_Nagelkerke_R_Square_in_Logit_Regression_analysis.

Schielzeth, H., N. J. Dingemanse, S. Nakagawa, D. F. Westneat, H. Allegue, C. Teplitsky, D. Réale, N. A. Dochtermann, L. Z. Garamszegi, and Y. G. Araya-Ajoy. 2020. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods in Ecology and Evolution 11:1141–1152. https://doi.org/10.1111/2041-210X.13434.

Straffi, P. 2016. On Screen Protractor. v0.5. https://sourceforge.net/projects/osprotractor. Accessed June 2023.

Symonds, M. R., and A. Moussalli. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behavioral Ecology and Sociobiology 65:13–21.

Tukey, J. 1949. Comparing individual means in the analysis of variance. Biometrics 5:99–114.

U.S. Fish and Wildlife Service. 2013. Eagle Conservation Plan Guidance Module 1 – Land-based Wind Energy, Technical Appendix D. Version 2. Division of Migratory Bird Management, Washington, D.C., USA.

https://www.fws.gov/guidance/sites/guidance/files/documents/eagleconservationplanguidance.pdf.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, NY, USA.

Wood, S. N. 2017. Generalized Additive Models: An Introduction with R. Second edition. CRC Press, Taylor & Francis Group, New York, NY, USA.

Zar, J. H. 1998. *Biostatistical Analysis*. Pearson.

Zuur, A. F., E. N. Ieno, and C. S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 1:3–14.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, New York, NY, USA.

Section 5. Technical Scope and Objectives

5.1 Budget Period 1: Develop a detailed, peer-reviewed study design for expanded study, and evaluate results of California pilot study (Tasks 1-4)

Timeline: 1 June 2017 – 30 November 2018 (Q1:M1 – Q6:M18)

Status: Completed

Objectives:

1. Develop a peer-reviewed study plan for 1) testing of the DTBird System at the Washington host facility, and 2) conducting integrative analyses of information gathered at multiple study sites.
2. Evaluate results of independently funded California pilot study to inform refinements of the DTBird system and form the basis for developing the Washington-based study plan.
3. Complete an expanded evaluation of false positives at the California wind facility.

Outcome Summary:

The project team completed a final peer-reviewed Study Design, including a set of proposed Quantitative Performance Targets (QPTs) DOE approved in October 2018 (Attachment 3). In August 2018, H. T. Harvey & Associates completed an expanded 10.5-month assessment of false-positive detections from the California pilot study at the Manzana Wind Power Project. They revealed that, of the video clips identified and categorized, 63% involved false positive detections and 61% of those events triggered a deterrent signal. Eagles represented 2% of these detections, but identifying targets to species was difficult based on low-resolution DTBird videos (DTBird does not automatically identify nor enumerate targets; technicians must do that by reviewing extracted video clips). In evaluations of the pilot study, recommendations suggested that Liquen focus on (1) adjusting the duration of the deterrent signal and signal

criteria, (2) applying AI to reduce false positives, (3) increasing the accuracy and precision of the spatial targeting to increase consistency of deterrents triggered by at-risk targets, (4) refining algorithms to enable target detections against landscape backdrops. The QPTs were established based on the pilot study, assuming the DTBird systems would meet or exceed the proposed performance targets at the Washington facility. In coordination with DOE and reviewers, the research team established a QPT range of 53–73% for the probability of detection. Greater than 50% was established as the QPT for successful deterrence of eagles. The false positive QPT established that <36% of all screened event records should involve targets determined to be false positives, including inanimate objects and non-target birds.

Budget Period 1 (Go/No-go) Outcomes(s) (Q6: M17-18):

During Budget Period 1, REWI completed all SOPO tasks and milestones. The final study design was submitted to DOE in July 2018, and the final response to remaining peer review comments was submitted to DOE in August 2018. Documentation of the recommended upgrades, the false positive report, and the QTPs were submitted to DOE in August 2018. REWI requested a six-month extension to Budget Period 1 due to delays in the Award Negotiation process, a U.S. Fish and Wildlife Service roadblock in the NEPA analysis for the initial study site resulting in a need to identify a new site for the 2-year experiment, and the process of peer reviewing the study design. An award modification was provided in August 2018 and included the following:

- DOE approved a six-month extension.
- Relocation of the 2-year experiment from original host site in Wyoming to facility in Washington State
- DOE granted the project additional funds for unanticipated work associated with the delays.
- An additional $200,000 in cost share for additional DTBird units used to increase the sample size of recordings of *in situ* raptor behavior responses and a more robust dataset.
- A revision of the study design to include a halfway checkpoint through the first year of the 2-year experiment whereby the project team analyzes the data accumulated and assesses whether there will be enough data in the first year to add a preliminary assessment of habituation to the study in the second year.

5.2 Budget Period 2: Expand evaluation of DTBird Detection and Deterrence Systems (Tasks 4-8)

Timeline: 1 December 2018 – 31 October 2022 (Q7-Q22:M19-M65)

Status: Completed

Objectives:

1. Install DTBird and complete the first year of a controlled experiment at the Washington facility designed to evaluate DTBird's effectiveness as an impact minimization technology.

2. Expand to a full year evaluation of DTBird deterrence capabilities at the California wind facility, focused on evaluating behavioral responses of *in situ* eagles based on DTBird video footage.
3. Conduct UAV flight trials at Washington wind facility.
4. Conduct a mid-year assessment to ascertain whether sufficient data were collected through Year 1 of a two-year controlled experiment to determine "proximate" effectiveness of DTBird with reasonable statistical power.

Outcome Summary:

Due to delays and equipment failures, of the 18 units originally proposed, only 14 DTBird units were installed in Washington, and usable data were provided by only 11 of these units during Year 1 of the Goodnoe Hills study, with no other departures from the approved study design. An estimated 53 ± 16.7% of confirmed eagles were considered to have been effectively deterred in an evaluation of the expanded dataset from the Manzana facility. The proportion of false negatives as determined from UAV flight trials at the Goodnoe Hills was 37 ± 10.7%, essentially identical to the estimate from the Manzana pilot study. In contrast, the rate at which false positives triggered deterrent signals substantially exceeded the relevant QPT at Goodnoe Hills. True false positives (TFPs; i.e., detections triggered by inanimate objects, insects, and software limitations/glitches) triggered an average of 3.6 ± 0.79 deterrent signals/turbine/day and resulted in an average of 1.9 ± 0.42 minutes of deterrent signaling/ turbine/day at Goodnoe Hills. Non-target avian false positives (NTAFPs; i.e., birds other than focal large soaring raptors) triggered an average of 2.2 ± 0.86 deterrent signals/turbine/day and resulted in an average of 1.2 ± 0.48 minutes of deterrent signaling/turbine/ day at Goodnoe Hills. When averaged across turbines, the probability of detection was 63 ± 11% at Goodnoe Hills, which was similar to the estimate derived from the previous Manzana study (63 ± 10%) and fell in the middle of the established QPT range (53–73%). The overall probability of detection estimate derived from combining data across all turbines and trial sessions at Goodnoe Hills (67%) also fell within the QPT range.

Budget Period 2 Go/No-go Outcomes (Q20 – Q22: M60-64):

During Budget Period 2, REWI completed all SOPO tasks and milestones. The project team requested a 12+ month extension of BP2, which the DOE granted in September 2020 to complete the commissioning of all DTBird units and the UAV flight trials. The project team came together in Spring 2021 to reschedule and rescope BP2 tasks, given the anticipated extension request. The project team completed two successful rounds of UAV flight trials to evaluate the detection and deterrent-triggering functions of DTBird at Goodnoe Hills in August 2021 and July 2022. The expanded full-year evaluation of *in situ* raptor behavioral responses to DTBird deterrents at the Manzana facility was reported on in August 2019. The project team completed the first year of the controlled experiment at the Goodnoe Hills facility in August 2022. This experiment evaluated the ability of DTBird to deter eagles and surrogate raptors from entering the RSZ of DTBird-equipped turbines. Based on the mid-year statistical-power assessment of data collected through Year 1, the project team recommended continuing a second year of the controlled experiment focused on evaluating DTBird's proximate

effectiveness, instead of pivoting to the alternative objective to evaluate potential habituation behavior. An updated budget justification was provided for BP1 and BP2 to include the following changes:

- An additional $200,000 in cost share for the overall project
- An additional 5 DTBird units (total 18) for the Washington experiment
- Addition of allowable indirect costs not included in the original budget justification.
- Increase in the budget for DTBird installation at the Washington site to reflect increased units.
- Redistribution of funds intended for trained raptor flight trials (originally task 6) to now cover additional screening and analysis of DTBird data (UAV flight trials and *in situ* raptor videos)
- Redistribution of funds allocated to REWI, Liquen, and H. T. Harvey & Associates to better reflect accurate predictions of project needs based on the completed pilot study and BP1.

5.3 Budget Period 3: Complete primary or alternative controlled experiment & video evaluation at Washington Facility; Conduct multi-site analyses (Task 8-12)

Timeline: 1 September 2022 – 31 May 2024 (Q22-Q28:M65-M84)

Status: Completed

Objectives:

1. Based on results of mid-year statistical power assessment, either extend to two years the controlled experiment focused on evaluating proximate effectiveness or pivot to Alternative Objective at Washington wind facility.
   1.a. Alternative Objective: Complete one (1) year of a controlled experiment at the Washington facility designed to assess the potential for habituation behavior.
2. Conduct multi-site analyses of field data.

Outcome Summary:

The project team completed the classification of in-situ raptor responses to deterrents. The resulting estimate of the proportion of successful deterrence responses with turbines spinning and deterrents broadcasting (53–100%) exceeded the established QPT of ≥50% successful deterrence for eagles. Additionally, the two-year experiment results indicated that broadcasted DTBird deterrents significantly reduced the dwell time of eagles around relevant turbines, especially when combined with elevated rates of deterrent triggering caused by false positives. However, broadcasted warning signals did not significantly influence the rate at which eagles triggered dissuasion signals, partly because eagles often entered the dissuasion signal zone without first being detected by DTBird within a warning signal zone. In the multi-site analysis, false positives were distinguished as TFPs and NTAFPs. Liquen adjusted the algorithms in January 2023, lowering TFPs from 3.9 to 0.8 triggers/turbine/day at spinning turbines. Post-

adjustments, the TFP triggering rate fell within or under the established QPT (1.6–2.8 triggers/turbine/day). Overall, turbine-specific counts of TFPs varied by site and 28-day sampling cycles.

Proportions of false negatives were determined by evaluating the number of UAV flight transects that should have triggered a DTBird detection but did not. The multi-site analysis revealed similar probabilities of detection at both sites (66% at Manzana and 64% at Goodnoe Hills) which exceeded the QPT established from the pilot study (≥63% detection probability or ≤37% false negative proportion).

The multi-site analysis of detection and deterrence triggering performance based on UAV flight and landscape characteristics revealed situation-specific landscape variations between Manzana and Goodnoe Hills that led to variability in DTBird's ability to detect and target objects of interest. Cloudy skies, wind speed, different UAV models (potentially reflecting differences in sexes and age classes), UAV speed, and pitch and roll angles all influenced the DTBird response distances.

The multi-species and golden eagle analyses confirmed significant differences in the probability of successful behavioral responses to broadcasted DTBird deterrents at the two study sites and indicated an effect of pre-exposure risk as well as an interacting effect of wind speed and raptor species. The probability of effective deterrence generally was highest for eagle/large raptors classified as at moderate risk of exposure to collision, likely because such birds had more time to respond effectively before entering the high-risk RSZ compared to birds that were initially at high risk of exposure.

Overall, for golden eagles flying within DTBird's calibrated maximum detection range for the species, the operation of DTBird appeared to reduce the likelihood of approaching the RSZ of spinning turbines by 20–30%. The study results also emphasized that DTBird's overall effectiveness may vary in different landscape/climatic settings and depending on the focal raptor species.

Section 6. Award and Modifications to Prime Award and the Statement of Project Objectives (SOPO)

Overall, the project award received fourteen modifications associated with delays and extensions, personnel changes, cost-share assistance, COVID-19, and technical DTBird system-related issues:

Modification 1 was created to allow for the deletion and replacement of Special Terms and Conditions to incorporate the following revisions: (a) delete and replace Term 13 Publications changing 'Wind Program' to 'Wind Energy Technologies Office' and (b) delete and replace Term 26 Cost Sharing to authorize providing the cost share on a Budget Period basis.

Modification 2 was created to delete and replace the SOPO and Federal Assistance Reporting Checklist and extend the period of performance from to 6/1/2017 – 8/21/2021 and adjust the special terms and conditions to delete and replace Term 8, NEPA requirements.

Modification 3 was created to delete and replace the Federal Assistance Reporting Checklist.

Modification 4 was created to obligate an additional $16,558 in Federal funding for the award, to delete and replace the SOPO, Federal Assistance Reporting Checklist, and Budget information, in addition to deleting and replace Term 26 Cost Sharing and Term 29 Indirect Costs.

Modification 5 was created to update the DOE Award Administrator.

Modification 6 was created to approve the continuation application and allow the recipient to move from Budget Period 1 to Budget Period 2. It also extended the period of performance for the award to 6/1/2017 – 5/31/2022, adding a 12+ month extension in Budget Period 2 and continuing to increase the cost share on the awardee's end.

Modification 7 was created to extend the period of performance to 6/1/2017 – 5/31/2023, revise the Government share, cost share, and total, provide additional funding, delete and replace the SOPO and Budget information, and delete and replace the Special Terms and Conditions to add Term 41, Foreign National Access Under DOE order 142.3A "Unclassified Foreign Visits and Assignments Program", Term 14 Publications, Term 26 Cost Sharing, and Term 32 Payment Procedures were also deleted and replaced.

Modification 8 was created to update the DOE Project Officer.

Modification 9 was created to extend the period of performance by 13+ months to 6/1/2017 – 5/31/2024 with Budget Period 2 specifically extended from 12/01/2018 – 09/30/2021 to 10/31/2022, and deleted and replace the following terms in the Special Terms and Conditions, Terms 41 Foreign National Access and add Term 42 Environmental, Safety and Heath Performance of Work at DOE Facilities, Term 43 Export Control, and Term 44 Prohibition on Certain Telecommunications and Video Surveillance Services or Equipment.

Modification 10 was created to delete and replace the SOPO, Budget Information, Term 26 Cost Sharing in the Special Terms and Conditions and reconfirm the project period. This modification confirmed the start and end date for the rest of the project timelines, per MOD 9.

Modification 11 was created to correct the period of performance start date.

Modification 12 was created to approve the continuation application, allowing the recipient to move from Budget Period 2 to Budget Period 3; approve the extension of the period of performance end date; update the recipient cost share and total project costs; and delete and replace Term 26 Cost Sharing and Term 29 Indirect Costs.

Modification 13 was created to update the DOE Award Administrator.

Modification 14 was created to update the Recipient Principal Investigator.

Section 7. Issues and Changes in Approach

During BP2 and towards BP3, due to delays in equipment shipping, personnel changes, and equipment challenges all exacerbated by COVID (Liquen staff could not travel to the U.S. to expedite addressing equipment issues), REWI requested and received multiple project extensions resulting in the project continuing three plus years after the originally proposed end date. Additionally, the project team contributed substantial added cost share to cover unanticipated costs related to the project challenges, as well as additional support provided by DOE related to COVID-19.

Equipment: 16 units were ordered, built, shipped, installed, and attempted to be commissioned. However, due to COVID restricting in-person servicing by Liquen, only 14 units were determined to be commissionable within a feasible timeline. Therefore, 16 units remain as Equipment costs, with Liquen covering the costs of the two unusable units as cost share.

Cost share changes:

- H. T. Harvey & Associates: additional cost share provided for UAV preparations.
- Liquen: additional cost share for two unusable units, as noted above.
- PacifiCorp: shifting BP allocation of cost share, as noted above. Additional costs to capture additional labor by PacifiCorp contractors in lieu of on-site support by Liquen (Liquen could not travel due to COVID).
- Portland General Electric: shifting cash cost share from external funders into BP2 to reflect anticipated timing of applying these funds.
- Puget Sound Energy: shifting cash cost share from external funders into BP2 to reflect anticipated timing of applying these funds.
- REWI: additional cost share in BP3 to ensure total project cost share percentage remains the same (57.76%).

During BP3, consistent issues occurred for DTBird systems at various turbines at the Goodnoe Hills sight, which required a considerable amount of additional project team time and PacifiCorp staff time to mitigate issues. Most notable were issues related to a delay in camera lens cover replacements, which were to be completed every six months and were not replaced until over a year after their initial replacement. Additional issues arose due to camera outages, analysis unit replacements, Vesta's server disconnections, and the unexpected Bonneville Power Administration (BPA) outage that lasted for approximately 26 days in May 2023. PacifiCorp was able to provide data that helped overcome a complete DTBird system communication failure after the BPA power outage that extended for the remainder of the study, but this resulted in Liquen and H. T. Harvey & Associates spending time beyond their scopes to format and align the PacifiCorp data and records stored in the DTBird on-line Data Analysis Platform (DAP).

Section 8. Task Accomplishments & Milestones

8.1 Task 1.0: Project Launch and Development of Peer-reviewed Study Design

8.1.1 Milestone 1.1.1: Completed peer-reviewed study design and quantitative performance targets (Q5:M15)

The final, peer-reviewed study design and a companion document, Response to Peer Reviewer Comments, were **submitted to DOE in July 2018 (Q5:M14)**. Following coordination with DOE, The National Renewable Energy Laboratory (NREL), and peer reviewers, the updated study design was **submitted to DOE in August 2018 (Q5:M15**) and **approved by DOE in October 2018 (Q6:M17).** In the original SOPO, the study design was intended to be completed in the third month of the project, but multiple delays pushed back the process, so the final draft was submitted to the DOE in the 14th month of the project. The study design was updated with respect to resident raptor observations at the Washington site to include a "partial year assessment" of data collected in the first half of Year 1 of the 2-year experiment to determine whether enough data were likely to be available to effectively analyze the "proximate" effectiveness of DTBird's audible deterrents for deterring eagles and other raptors. The additional assessment was designed to determine if the DTBird units would be assigned a continued daily control-treatment rotation schedule in Year 2, designed to minimize the potential for turbine-specific habituation, or a permanent control or treatment mode during Year 2 to enable evaluating the potential for habituation.

8.2 Task 2.0: Evaluate false positives using data collected during the pilot study at California wind facility

8.2.1 Milestone 2.1 False positive rates quantified at California wind facility (Q5:Q15)

A final report on false positive detections at the Manzana Wind Power Project in southern California was **submitted to DOE in August 2018 (Q5:M15)**. H. T. Harvey & Associates analyzed footage collected by DTBird from December 2016 through October 2017. The team sampled 5,212 detection events and were able to classify the detected targets in 5,208 of those records. The classifications included estimates of 33% (1,712) TFPs 30% (1,567) NTAFPs. There was an average of 2.2 TFP detections/turbine/day, and 1.9 NTAFP detections/turbine/day. Of the TFP detections, 61% triggered a deterrent signal, resulting in an average of 1.7 extraneous deterrent signals/turbine/day. Of the NTAFP detections, 58% triggered a deterrent signal, resulting in an average of 1.2 deterrent signals/turbine/day. False positive rates varied among turbines and months. For example, there was an increase in TFPs caused by insects in June reflecting a seasonal increase in insect abundance, rather than variation in DTBird performance. Of the TFPs, 45% were associated with various aircraft, 23% turbine blades, 21% various sky artifacts, 8% insects, 1% precipitation, and <1% other objects such as balloons or floating leaves. Eagles

represented approximately 2% of all detections, however, DTBird does not identify or filter targets by species. Per their preceding standard practice, Liquen implemented some adjustments to the filtering algorithms that significantly reduced the instances of TFPs triggered by turbine blades after February 2017, more than 2 months after the systems were declared "fully commissioned."

8.3 Task 3.0: Evaluation of pilot study

8.3.1 Milestone 3.1 Recommended updates to DTBird system delivered to technology vendor (Q6:M16)

With insights from the DTBird 2016–2017 pilot study at the Manzana Wind Project, H. T. Harvey & Associates, Liquen, and REWI created a set of feasible recommendations for future updates and upgrades to the DTBird system. **These recommendations were submitted to DOE in August 2018 (Q5:M15)** and **received review from DOE in September 2018 (Q6:M16)**. All comments and revisions to the recommended system updates **were submitted to DOE in December 2019 (Q7:M19).** REWI and Liquen recommended updates focused on the following four topics:

1. Reducing the duration of deterrent signals and signal criteria to ignore fast moving targets that cannot be birds.
2. Artificial intelligence/machine learning capabilities to reduce false positives.
3. Increasing accuracy and precision of spatial targeting to increase the consistency of deterrents signaled in response to at-risk targets.
4. Refine algorithms to enable target detections against landscape backdrops.

8.3.2 Milestone 3.2 QTPs established based on analysis of pilot study (Q6:M18)

A set of QPTs were ***proposed to DOE in the study design submitted August 2018 (Q5:M15)***. Performance targets were established based on the results of the Manzana pilot study and with the expectation that DTBird would meet or exceed the performance targets.

As part of Milestone 3.2, the following QPTs were established for future Milestones 6.1-7.1, to be completed in BP2:

Milestone 6.1: 53-73% overall UAV detection rate; false negative rate (inverse of detection rate) 27-47%. This detection rate was selected based on data from the pilot study, in which 63% ± 10% SD of the UAV flights were detected in the flight trials. The false negative rate (inverse of the detection rate - UAV flights that occurred but were not recorded by DTBird), was 37% ± 10% SD.

Milestone 6.2: ≥ 50% successful deterrence rate for eagles. At least 50% of bald and golden eagles should exhibit avoidance behavior to a DTBird system within 5 seconds of the deterrent signal when sound is "on". This deterrence rate was selected based on data from the pilot study, in which 36% of raptors (overall) responded effectively to the deterrent signals. We selected this target because our expectation is that DTBird's performance in the two-year study

will meet or exceed the performance observed in the pilot study. The project team anticipated having access to higher resolution video footage for the DOE study, and the ability to better identify birds in the DTBird video clips.

Milestone 6.3: Not to exceed 1.6 -2.8 TFPs triggers/turbine/day or ≤36% of total video records collected by DTBird units. This false positive rate is based on data from the False Positives Analysis of data collected during the pilot study, in which DTBird systems detected 2.2 ± 0.64 TFPs/turbine/day, and 36% of video records from DTBird contained targets determined to be non-avian objects (e.g., turbine blades, aircraft, insects, raindrops).

Milestone 7.1: Complete a summary of the mid-year assessment and provide a recommendation of which objective to pursue in Year 2 of the experiment (proximate effectiveness or habituation focused).

The team provided responses to comments regarding the proposed target detection rates of 53% for eagle-surrogate UAVs ***in September 2018 (Q6:M16)***. As part of this response, the team clarified that a lower confidence interval would be incorporated for drones (63% ±10% detection rate) and accommodating unidentified large birds for the eagle detection rate (36% detection rate) due to the likelihood that an unknown percentage of unidentified large birds were eagles. Since the DTBird system does not distinguish between eagles and other bird species that trigger detections and deterrents and in the United States, where the primary species of concern are bald and/or golden eagles, triggering for non-target species could be deemed excessive. Therefore, the project team sought to quantify the percentage of each event type triggered by eagles (i.e., detections, warning signals, and dissuasion signals) compared to non-eagles and false positives as well as the rate of deterrent signal triggers per turbines per day.

8.4 Task 4.0: Update DTBird system and revise study design for BP2 and BP3 as appropriate

8.4.1 Milestone 4.1 Study design revised (Q7:M19)

A revision to the study design was **submitted to DOE in August 2018 (Q5:Q15)**, response to comments from expert review were **submitted to DOE in January 2019 (Q7:M20)**. The following was clarified/updated:

- The presumption of an effective DTBird deterrent signal relative to North American eagle species, as Liquen's rational noted that the deterrent signal was developed for a range of species both European and North American (including golden eagles). The presumed similar reaction to the technology is expected by bald eagles and European white-tailed eagles, both in *Haliaeetus*.
- The rational for using the Akaike Information Criterion (AIC) in the proposed statistical approach. H. T. Harvey & Associates noted that the AIC had several potential predictor variables under consideration. AIC scores would be used to support deriving an optimized model best predicting future observed outcomes while minimizing predictor variables. Additionally, use of AIC was with respect to the detection efficacy related to drones, not eagle activity or deterrence.

- An evaluation of the pilot study was completed with the adjusted false positive and deterrent response analyses of the California dataset; the results of the two analyses informed the decision to not revise the study design.
- In addition, the numbers of units scheduled for deployment at the Goodnoe Hills site was adjusted from 18 to 16; concerns were raised about whether this change would compromise the study design or project objectives, but ultimately determined it would not affect the statistical power of the study.

8.4.2 Milestone 4.2 Updates to DTBird system completed (Q8:M22)

The recommended DTBird system updates were **submitted to DOE in August 2018 (Q5:M15)** and were **provided to Liquen in November 2018 (Q6:M18)**. Liquen confirmed the incorporation of the following updates in advance of the next phase of fieldwork at the Goodnoe Hills site.

- An increase in the size of the red box in the DTBird video footage showing the location of targets.
- An increase in the size of the time stamps on video footage.
- DAP ID number included in file name by default.
- Use of higher megapixel resolution to facilitate improved target identification.
- A second ring of speakers to broadcast deterrent signals.
- Increased frequency of replacement of lens covers to reduce impacts of sun glare, specifically polarized lenses, if feasible.

All updates were incorporated with one exception regarding a change to polarized lens covers which needed further assessment, therefore the project team decided to modify preventative maintenance of camera lens cover changes from annual to biannual.

Longer term recommendations for updates that were to continue to be researched and developed include:

- A new statistical analysis will be done by DTBird, and results incorporated in the second year of fieldwork to optimize the sounding of the deterrence trigger and refine the triggering criteria and duration of signal.
- Liquen's evaluation of whether a sample turbine at Goodnoe Hills can incorporate a set of polarized lenses during Year 1 data collection compared to standard lens but has not determined whether this would complicate the experimental design.
- An updated version of the DTBird software set to be implemented at Goodnoe Hills, which will allow the operation of AI software over the detections of the videos, data collection and preliminary testing set to be completed following the first year of service.
- Improvements to the precision of detection, warning-trigger, and dissuasion-trigger targeting to ensure effective deterrence responses and reducing unnecessary signaling are expected to be longer-term, but increased image quality, new software algorithms, and sound triggering criteria, as noted above, should reduce unnecessary deterrence signaling.

- Incorporation of an algorithm that summarizes all instances of video failure at individual cameras to allow more comprehensive evaluations of system performance, is set for Year 2 of field testing at Goodnoe Hills.

All 16 DTBird units arrived at Goodnoe Hills, all but four were partially installed due to special equipment required to install a second ring of speakers. Liquen reverted to the previous software platform due to technical difficulties, but the previously noted recommendations were included prior to UAV flight trials and Year 1 data collection **completing this milestone in May 2021 (Q16:M48)**.

8.5 Task 5.0: Install DTBird systems at Washington wind facility

8.5.1 Milestone 5.1 DTBird systems installed and commissioned at Washington wind facility (Q13:M37)

Liquen and PacifiCorp experienced several unanticipated delays in acquiring parts for units, sending personnel for installation, and integrating the DTBird units to the SCADA and network amid the COVID-19 pandemic. All 16 DTBird units **arrived at Goodnoe Hills in October 2019 (Q11:M29)** where all but four, due to weather concerns, were assembled and fully installed. PacifiCorp contracted a Vestas crew on site to conduct their part of the cost-share to conduct assembly and installation. Liquen and Pacificorp began integration of the units into the Washington wind facility's online network. **In the Spring of 2020 (Q12)** Liquen and PacifiCorp continued integrating the 12 fully installed units into the online network and enabled remote control access. **A 13+ month extension was awarded in September 2020 (Q14:M40)** starting with BP2 to allow for the complete commissioning of DTBird units and to allow for UAV flight trials. Thirteen of 14 DTBird units were fully commissioned for the Year 1 experiment, however only 11 were operating sufficiently and consistently enough to yield usable data. Following a myriad of system maintenance issues involving camera outages, communication failures, and analysis units without remote access, all 14 DTBird units **were fully commissioned and effectively functional in early September 2022 (Q22:M64)**, though additional significant (> 30 days) gaps in functionality continued for several units.

8.6 Task 6.0: Expand evaluation of *in situ* bird video footage at California and Washington wind facilities, conduct UAV flight trials at Washington wind facility, and analyze site-specific results

8.6.1 Milestone 6.1: UAV flight trials completed at Washington wind facility (Q18:M53)

H. T. Harvey & Associates and Remote Intelligence attempted to conduct an initial round of UAV flight trials at the Washington facility during May 2021 (Q16:M48); however, excessive wind and a DTBird system failure resulted in no usable data from that attempt. **The first round of successful flight trials at this facility occurred in early August 2021 (Q17:M51).** This trial session involved two UAV models and provided approximately 8 hours of usable data collected

at three DTBird turbines, but ended prematurely when both UAVs were destroyed in crashes. In September 2021, H. T. Harvey & Associates coordinated with Remote Intelligence to prepare two new UAV aircraft– for use during a third round of flight trials. The third round was then conducted during the last week of July 2022 and provided additional data collected at four DTBird turbines, but with one of the new aircraft also destroyed in a crash caused by an equipment failure. UAV flight trials at Goodnoe Hills were **completed at the end of July 2022 (Q21:M62)**, with useable data collected during 29 individual, automated flight missions conducted at five DTBird-equipped turbines, which yielded 482 distinct flight-transect samples suitable for analysis. Results of this site-specific investigation were provided as part of the **project's Continuation Application in September 2022 (Q22:M64)**.

8.6.2 Milestone 6.2 DTBird video data collection and enhanced site-specific evaluation of *in situ* eagle responses to deterrents completed for California facility, with evaluation restricted to first year of data collection (Q10:M28)

The objective was to expand a preliminary pilot-study evaluation based on one full year of data collected at the Manzana wind facility to support quantification of the effectiveness of DTBird audible deterrents to deter golden eagles from approaching equipped turbines. The analysis evaluated the behavioral responses of *in-situ* eagles to deterrence signals using DTBird data collected from January through December 2017 at the seven turbines outfitted with DTBird systems. Event data recorded in the DAP were processed and analyzed to estimate the probability of effective deterrence for golden eagles and other large soaring raptors.

For this analysis, H. T. Harvey & Associates randomly selected 10 days/month during the sampling period and collected behavioral data for all large raptors detected by the DTBird system and exposed to deterrent signals. The following behaviors were recorded for each relevant deterrence event: approximate direction of travel relative to turbine before and after signal emittance, risk of approaching RSZ before signal emittance (i.e., *Preexposure Risk)*, whether the raptor appeared to respond to the deterrent signal, and whether the raptor's response to the signal reduced its risk of approaching the RSZ (i.e., *Reduced Risk*). H. T. Harvey & Associates used this final behavioral classification, *Reduced Risk*, as the response variable in a series of general linear models (GLMs) to evaluate the influence of month, initial risk level (*Risk*), and raptor group (i.e., eagles, vultures, and buteos) on the probability of deterrence.

The assessment indicated that, across all evaluated events, the DTBird collision avoidance module effectively deterred at least 53% confirmed golden eagles, 57% of turkey vultures, 38% of buteos, 64% of falcons, and 43% of all raptors combined. Adding in cases where deterrent responses were classified as potentially effective elevated the estimated probability of deterrence for golden eagles to 74%, for turkey vultures to 81%, for buteos to 69%, for falcons to 100%, and for all raptors combined to 72%.

The GLM results suggested that birds at moderate to high risk of approaching the RSZ were more likely to respond effectively and tended to divert away from the risk zone more strongly when exposed to the deterrent signals compared to birds that were at low risk of exposure. The GLMs also suggested that eagles were slightly more likely than buteos to respond to deterrence

signals and responded with greater diversion angles. The interpretability of these results had limited rigor due to modest sample sizes and the lack of control (i.e., to support evaluating the differential effect of spinning turbines alone versus spinning turbines plus broadcasted DTBird deterrents). H. T. Harvey & Associates **produced a final summary report in August 2019 (Q9:M27)** documenting the enhanced analysis.

8.6.3 Milestone 6.3 Preliminary site-specific estimates of rates of false positives and false negatives produced for Washington wind facility (Q20:M60)

The probability of false negatives as determined from UAV flight trials conducted at Goodnoe Hills was 37 ± 10.7%, identical to the false negative probability estimated for the Manzana site based on the pilot study (37 ± 10%). This outcome suggests consistent performance of the primary detection functions of the DTBird system.

A probability of detection (converse of probability of false negatives) logistic GLM (LGLM) analysis provided additional perspective concerning factors that influence the overall probability of DTBird detecting an eagle-like UAV if it flies anywhere through the 240-meter-radius detection envelope projected based on calibration for golden eagles. The limitation of this analysis was that for flights that were not detected (false negatives) there were no reference points to use for precisely characterizing the flight, location, and environmental characteristics at the time of a specific event to use as covariates. Consequently, we focused attention on discerning the influences of only a select few metrics derived by averaging across all points along a given sample flight. Nevertheless, this approach illustrated some variability in the probability of detection through the day, perhaps related to influence of the sun's position and intensity, but more importantly emphasized that the probability of detection was highest when the target flew at moderate distances from the turbine (generally high with average flight distances of 80–160 m) through the mid-section of the camera viewshed (generally high with viewing angles from camera up to UAV of 25–40°). The probability of detection averaged lower when the target flew closer to or farther away from the camera or within the lower or upper margins of the camera viewshed.

The rate at which false positives triggered deterrent signals exceeded the QPT during Year 1 of the Washington study. TFPs triggered an average 3.6 ± 0.79 deterrent signals/turbine/day and resulted in an average of 1.9 ± 0.42 minutes of deterrent signaling/turbine/day. NTAFPs, defined as birds other than focal large raptors (eagles, vultures, buteos, and ospreys), averaged 2.2 ± 0.86 deterrent signals/turbine/day, and 1.2 ± 0.48 minutes of deterrent signaling/turbine/day. Calculation of False Positive rates was based on 12,962 detections (defined as triggering a video recording). An estimated 66% of those detections resulted from TFPs, 25% from NTAFPs, and 9% from presumed eagles and other large raptors. Approximately 33% of the TFPs and 52% of the NTAFPs triggered a deterrent signal. TFPs averaged 11.2 ± 3.46 (SD) detections/turbine/day, and NTAFPs averaged 4.2 ± 1.49 detections/turbine/day. A report was produced to accompany this milestone in the **Continuation Application submitted in September 2022 (Q22:M64)**.

8.6.4 Milestone 6.4 Initial site-specific models developed to quantify the spatial accuracy of the DTBird detection and deterrent-triggering system at Washington wind facility (Q20:M60)

UAV flight-trial data were analyzed, and site-specific statistical models were developed to evaluate the influence of flight and environmental covariates on DTBird detection and deterrent-triggering response distances at the Goodnoe Hills facility. This site-specific analysis used the final Manzana site-specific model as a template but did not consider the same full suite of variables initially considered to develop the Manzana model. We made this choice primarily because unanticipated and unavoidable lengthy delays in our ability to conduct a successful series of UAV flight trials at the Goodnoe Hills severely compressed the time available for analysis ahead of the deadline for submitting a BP2 Continuation Application to the DOE to support moving into BP3. Nevertheless, this initial modeling effort was informative in suggesting similar relationships at the two sites for several key covariates of interest. We then developed further insight during BP3 in conducting an integrated assessment based on combining data from the two sites (see Task 10.1).

The initial Goodnoe Hills response-distance modeling effort revealed both similarities and some notable differences compared to the previous Manzana modeling effort. The similarities included (a) different UAV models influenced the response distances similar to what one might expect to occur in relation to eagles of different sizes and colorations, (b) eagle-like UAVs flying at relatively high altitudes were detected at greater distances than those flying at lower altitudes relative to the turbine base, c) greater degrees of ascent and rolling from side to side generally increased the response distances as a result of increasing the degree the UAV profile was exposed to the DTBird cameras, and (d) response distances were generally higher under uniformly bright mostly cloudy skies and lower under uniformly dark overcast skies, with clear blue and partly cloudy skies showing variably intermediate responses. Notable contrasts in the two sets of site-specific results included significant contributions of Turbine ID, solar intensity, wind speed, and UAV elevation angle in the Manzana model but not in the Goodnoe Hills model.

A report was produced to accompany this milestone in the **Continuation Application submitted in September 2022 (Q22:M64).**

8.7 Task 7.0: Conduct first year of controlled experiment at Washington Wind Facility

8.7.1 Milestone 7.1 First year data collection completed for controlled experiment (Q21:M63)

Thirteen of 14 DTBird units installed on turbines were fully commissioned, and H. T. Harvey & Associates and Liquen implemented the full 14-turbine 28-day muting rotation schedule for all available DTBird units. **The project team ended Year 1 experimental data collection in August 2022 (Q21:M63).**

8.7.2 Milestone 7.2 A summary of progress and findings up to 10 months of data will be included in the Go/No Go report for BP2. This summary will include a power analysis performed on at least 6 months of data, as well as a recommendation of which objective (proximate effectiveness or habituation) to pursue in Year 2 of the experiment (Q21:M63)

The estimated effect size based on the existing dataset of seven 28-day cycles was a 66% reduction in the number of dissuasion signals at treatment turbines compared to control turbines. This difference was not significant ($p = 0.192$); our statistical power to detect an effect of that amount with $\alpha = 0.05$ was estimated to be approximately 23%. The simulation portraying what we could expect based on an expanded 13-cycle dataset with comparable per-cycle sample sizes boosted the estimate of power to detect a 66% effect to approximately 42%, but still falling well short of an optimal power target of 80%. The 13-cycle projections suggested that we would have 80% power to detect only a >85% effect size, assuming the sampling results remained comparable to the initial 7 cycles across the extended 13-cycle period.

Based on these results, it was recommended to continue with the experiment as designed for testing DTBird's proximate effectiveness. **REWI submitted this mid-year assessment and recommendation along with its Continuation Application at the beginning of September 2022 (Q22:M64)**.

8.8 Task 8.0: Complete controlled experiment and analyze results

8.8.1 Milestone 8.1 First two months of controlled experiment's Year 2 DTBird data collected at Washington site (Q21:M65)

REWI, H. T. Harvey & Associates, PacifiCorp, and Liquen continued with the second year of the experimental design to ensure sufficient data to evaluate the proximate effectiveness of DTBird, starting in September 2022 (Q22:M64). Liquen implemented updates to the algorithm from November 2022 – March 2023, effectively allowing for full commissioning of the DTBird units for use in Year 2 of the experiment.

8.8.2 Milestone 8.2 Controlled experiment completed, and results analyzed; an estimate of eagle collision risk reduction from DTBird calculated (Q26:M78)

The 14 turbines included in the experiment were randomly assigned to control and treatment groups daily. Those assigned to the treatment group broadcasted warning and dissuasion deterrents when triggered. Those assigned to the control group triggered deterrents virtually but did not result in audible deterrents. Results indicated that the presence of broadcasted warning signals did not significantly influence the rate at which eagles triggered dissuasion signals, likely because eagles often entered the dissuasion signal zone without ever being detected by DTBird within the warning signal zone. However, results also indicated that broadcasted deterrent signals significantly reduced the time eagles spent near DTBird-equipped turbines

(aka dwell time). There was also a strong interactive effect of treatment group and false positive rates, meaning that if warning/dissuasion signals were triggered and broadcasted more frequently at treatment turbines by false positives, this led eagles to spend less time around those turbines. These results suggest that, despite our concerns that high false positive rates might cause eagles to become less responsive to deterrent signals, negative habituation did not occur over the 2-year course of this experiment. Instead, positive habituation appeared to occur over the 2-year course of this study, likely because the deterrent triggering rate caused by false positives was atypically high through the first 18 months of the study and the overall deterrent triggering rate was excessive during the last 5 months of the study due to a DTBird operations issue stemming from a facility-wide power outage. **A report for this milestone was submitted to DOE in December 2023. (Attachment 4).**

8.9 Task 9.0: Evaluate behavioral responses of raptors exposed to deterrent signals at Washington wind facility

8.9.1 Milestone 9.1 All DTBird video evaluation and classification of in-situ raptor responses to deterrent signals completed. Target performance is ≥50% successful deterrence for eagles (Q27:M79)

H. T. Harvey & Associates reviewed detection videos from the 14 DTBird turbines, sampling 10 randomly selected days within each 28-day period for the first year of the experiment. For all screened records in which a deterrent signal was triggered by a confirmed or probable eagle, vulture, or buteo, investigators evaluated the bird's flight behavior, including path divergence and changes in flapping style to classify each event into one of four response categories: Yes (Confirmed Effective), Potential (Potentially Effective), No Response, Not Relevant (response did not reduce risk). Records were also categorized by collision risk prior to deterrent exposure. H. T. Harvey & Associates then evaluated the differences in the categorical response proportions among the control (deterrents muted) and treatment (deterrents broadcasting) groups using a 2-way Pearson chi-square analysis and estimated the probable successful deterrence rate as the combination of responses classified as confirmed and potentially effective.

Of the 19 instances in which a golden eagle triggered a warning signal, 13 resulted in a successful or potentially successful deterrence response (68%). Similarly, 10 out of 19 of the instances in which a golden eagle triggered a dissuasion signal resulted in a successful or potentially successful deterrence response (53%). Small sample sizes did not allow for a comparative analysis between control and treatment groups for each species of eagle; however, based on the comparative control-treatment results and for all analyzed groups and species, broadcasted deterrents consistently resulted in at least a doubling of the proportion of cases where a successful or potentially successful response was evident. In addition, although the percentage of golden eagle responses classified as confirmed effective (32%) fell below the QPT established based on the Manzana pilot study, the combination of confirmed effective and potentially effective responses exceeded 50% for all analyzed species groups and for both warning and dissuasion signals (overall range 53–100%). **A report for this milestone was submitted to DOE in June 2023 (Q25:M73) (Attachment 5).**

8.10 Task 10.0: Complete combined multi-site analyses

8.10.1 Milestone 10.1 Multi-site analyses of detection and deterrence triggering capabilities as a function of flight and landscape characteristics completed (Q24: M72)

H. T. Harvey & Associates modeled DTBird response distances as a function of several environmental and UAV variables. The general linear mixed model (GLMM) that best explained variation in DTBird response distances suggested that response distances were more variable among the seven Manzana DTBird turbines than among the five Goodnoe Hills DTBird turbines where UAV flight trials were conducted, likely reflecting situation-specific landscape variation leading to modest variability in DTBird's ability to detect and target objects of interest. The average response distance at the Manzana facility was marginally lower than that at the Goodnoe Hills facility. Overcast skies significantly increased detection and deterrence-triggering distances compared to fair skies, suggesting greater detectability given a contrasting sky backdrop. Response distances tended to increase as the wind speed increased, but this relationship was only moderately significant. Increased wind speeds increased the degree to which the UAVs bounced around in the air, increasing profile exposure to the DTBird cameras and thereby increasing detectability as suggested by increasing response distances. Detection and deterrent-triggering response distances also tended to increase with UAV ground speed, suggesting that DTBird could more easily detect faster moving targets. DTBird response distances were also dependent on the interactive influences of UAV pitch and roll angles, whereby pitching and rolling acted in concert to increase exposure of the UAV profile to the cameras.

Output for the best performing GLMM also indicated that the different UAV models used accounted for a noteworthy difference in detection and deterrent-triggering distances. The two UAV models used at the Manzana site showed the greatest variance in response distances, whereas variation among the three distinct UAV models used at Goodnoe Hills was less pronounced. As the initial pilot study at Manzana suggested, response distances tended to be shorter for the skinnier-bodied UAV models when compared with the more eagle-like, robust-bodied models. Some of the variation observed potentially mimicked the kind of variation that could be expected with physical differences associated with eagles of different sexes and age classes.

We also note here that the differences rated as statistically significant effects given our data sometimes amounted to magnitudes that may not have especially noteworthy biological or operational significance (e.g., 10–20 m differences in detection range for birds that may easily move farther than that in less than a second). However, our study was not designed to specifically quantify the relative effectiveness of different calibrated detection and deterrent triggering distance thresholds nor the spatiotemporal aspects of what an eagle requires as deterrent warning to avoid calamity under different flight conditions. Therefore, we have no basis for presuming what may be biologically/operationally significant in this context.

The best performing GLMM did not include the UAV's climb rate, location of the UAV in relation to focal turbine (i.e., direction from turbine), the UAV direction of travel (i.e., course over ground), sun azimuth, solar irradiation, or sun elevation angle. **A report for this milestone was submitted to DOE on May 2023 (Q24:M72) (Attachment 6).**

8.10.2 Milestone 10.2 Complete multi-site analyses of false positives and false negatives (Q25:M73)

We used flight trials involving UAVs designed to coarsely mimic the general size, weight, and coloration of golden eagles to quantify the probability of false negatives (or conversely the probability of detection) at the two study sites. To investigate false positives, we used DTBird event data recorded for *in situ* eagles and other large raptors. A high false-positive rate during Year 1 of the Goodnoe Hills study proved concerning and required, upon DOE approval, that Liquen update the DTBird filtering algorithms to reduce the potential for blade-related, insect, and sky-artifact sources of false positives.

Based on the Manzana pilot study, a QPT for the probability of detection of 63% (or 37% false negatives) was established to evaluate the comparative performance of DTBird systems installed at the Goodnoe Hills. A multi-site GLMM revealed an overall 65% probability of detecting an eagle-like UAV within 240 meters or less of the cameras, with a nominally higher detection probability at the Manzana site (66%) than at the Goodnoe Hills site (64%). This outcome suggested consistent performance of the primary DTBird detection function at both sites. The results also indicated a higher chance of detection when the target flew within 80–160 meters of the turbine, versus closer or farther away, and at elevation angles that placed it within the middle of the camera viewsheds, versus high or lower.

The false positive analyses distinguished between TFPs (representing non-avian factors such as aircraft, insects, spinning turbine blades, and high-contrast sky conditions that sometimes trigger false detections, called sky artifacts) and NTAFPs (representing detections of birds other than focal large raptors, defined in this study as eagles, vultures, buteos, and ospreys). The established QPTs stipulated that (a) the overall TFP deterrent-trigger rate should not exceed 1.6–2.8 triggers/turbine/day; and (b) no more than 36% of all relevant and classified detections recorded by the DTBird systems should result from TFPs.

The TFP deterrent-triggering event rate exhibited interactive effects of site and time (28-day cycle). The overall average TFP rate at the Goodnoe Hills across the full 23 months of sampling was 3.9 TFP detections with deterrent triggers/turbine/day, which exceeded the established QPT. However, after Liquen made additional adjustments to reduce the false positive rate in January 2023, the TFP deterrent-trigger rate for the subsequent 7 months dropped to an average of 0.8 triggers/turbine/day, well below the performance target and comparable to contemporaneous rates at Manzana. Similarly, although TFPs resulted in more than 50% of all detections that triggered deterrents before the adjustments were made, the proportion dropped to 25% post-adjustments, again falling below the established performance target. Moreover, in both cases the post-adjustment rates at Goodnoe Hills were lower than at Manzana, suggesting

improvement in the filtering algorithms. **A report for this milestone was submitted to DOE in November 2023 (Q26:M78) (Attachment 7).**

8.10.3 Milestone 10.3 Complete multi-site analyses of behavioral responses of in-situ raptors to deterrence signals (Q26:M76)

To accomplish this task, H. T. Harvey & Associates observed video data of eagles and other large raptors approaching DTBird-equipped turbines and classified the eagles' behavior during/after deterrent signals were broadcasted. They included data from both the observational study at the Manzana Wind Power Project and Year 1 of the experimental study at the Goodnoe Hills Wind Farm. The analyses focused on evaluating individual eagle/large raptor behavior at both study sites in response to spinning blades with broadcast deterrents, but also qualified those multi-site results based on the site-specific control-treatment results from the Goodnoe Hills that allowed for quantifying the differential deterrence effects of spinning turbines alone versus spinning turbines plus broadcasted DTBird audio deterrents.

H. T. Harvey & Associates reviewed detection records and videos collected for one full year at both study sites, including at all seven DTBird-equipped turbines at the Manzana facility and 11 acceptably functional DTBird turbines at the Goodnoe Hills facility. Sampling at both sites included screening all records that triggered a deterrent signal on 10 randomly selected days/28-day period during the first year of data collection at each site. For all records where the target was classified as a confirmed or probable eagle, vulture, or buteo, investigators evaluated the bird's flight behavior, including path divergence and changes in flapping style, and classified its response to the triggered deterrent into one of four categories: Yes (Confirmed Effective), Potential (Potentially Effective), No Response, Not Relevant (evident response but did not reduce risk). H. T. Harvey & Associates then evaluated differences in the proportions of categorical responses by site using 2-way Pearson chi-square analyses for each raptor group (eagles, vultures, and buteos). Records were also classified and evaluated by Preexposure Risk, but limited sample sizes precluded preparing 3-way chi-square analyses including Site and Preexposure Risk as predictors.

H. T. Harvey & Associates used a LGLM to evaluate how the probability of effective deterrence (reduced to binary dependent variable) was influenced by Site, Species Group, Preexposure Risk, and Wind Speed. The multi-species and golden eagle specific LGLM analyses confirmed effects of Site and Preexposure Risk, and an interactive effect of Wind Speed and Species Group. The probability of effective deterrence was overall slightly higher at the Manzana site, for unknown reasons, but possibly reflecting factors such as region-specific species sensitivities, habituation patterns of resident vs non-resident birds, and variation in landscape or climatic features. The probability of effective deterrence was highest for birds classified as at moderate Preexposure Risk, likely reflecting such birds having more time than birds at higher risk to effectively respond before closely approaching the RSZ. Higher wind speeds resulted in a higher probability of effective deterrence for eagles and vultures, but not buteos, potentially because larger raptors are more reliant on, and capable of using, wind for in-flight maneuvering.

Because of the poor quality of videos that DTBird uses, we took a conservative approach to classifying behavioral responses, meaning that we categorized responses based on the degree to which we could discern a behavior as effectively reducing risk for the eagle. Partially in order to maintain a sufficient sample size for analysis, we chose to consider effectiveness estimates that included both confirmed effective and potentially effective deterrence responses to evaluate the potential for DTBird deterrents to reduce the risk of eagles and other large raptors entering the RSZ of spinning turbines at the two study sites. This may overestimate DTBird's effectiveness to some degree. However, it is just as likely, if not more likely, limiting the results to confirmed effective responses would have underestimated the rate at which DTBird effectively reduced risk for eagles, because of our limited ability to confidently discern and classify relatively subtle but nonetheless effective behavioral responses. All further results summarized below are based on statistics representing the combination of confirmed and potentially effective responses as the basis for estimating the probability of effective deterrence. **A report for this milestone was submitted to DOE in September 2023 (Q26:M76) (Attachment 8).**

8.10.4 Milestone 10.4 Produce a multi-site estimate of collision risk reduction, estimate of eagle fatality reduction (# eagles/year) attributable to DTBird completed (Q27:M79)

The objective was to estimate DTBird's overall effectiveness in reducing the risk of eagles entering the RSZ of spinning turbines, based on multiple complementary approaches. The first approach involved combining probability of detection estimates derived from the UAV flight trials with probability of effective deterrence estimates derived from the behavioral analyses. The multiplicative combination of these estimates yielded an estimated 52% reduction in the probability of confirmed golden eagles entering the RSZ of spinning turbines with broadcasted deterrents at the Manzana facility, and a 38% reduction at the Goodnoe Hills facility. Data for all eagles combined from Goodnoe Hills (rare occurrences of bald eagles at Manzana) revealed similar results for golden eagles alone, except limited data suggested effective deterrence was higher for bald eagles than for golden eagles.

The Goodnoe Hills control-treatment experimental setup confirmed the addition of DTBird audio deterrents increased the likelihood of effective deterrence compared to just spinning turbines alone with deterrent signals muted. Recalculating the estimates of detection and deterrence effectiveness for golden eagles alone based on the Goodnoe Hills control-treatment results yielded a 24% probability of DTBird audio deterrents reducing risk of entering the RSZ of spinning turbines at Manzana and 19% for Goodnoe Hills. Narrowing the focus further to estimating DTBird's effectiveness when an eagle-surrogate UAV was flying in core exposure locations and *in situ* eagles were classified for behavioral analysis as at moderate to high Preexposure Risk revealed that spinning turbines plus deterrents resulted in a 68% probability of reduced risk, with the added effectiveness of deterrents alone reducing estimated risk by 37%.

The second approach used to estimate risk reduction from DTBird was based on the Goodnoe Hills 2-year control-treatment experiment involving randomized daily rotations of muted and broadcasted deterrents. For golden eagles alone, the dissuasion-trigger (dependent variable =

probability of triggering a dissuasion signal) and dwell-time (dependent variable = eagle dwell time as reflected in extent of video recording) models yielded similar estimated reductions (27–29%) in the two dependent variables at DTBird-equipped turbines when the audio deterrents were broadcasted compared to when the deterrents were muted. Combining insight from both approaches suggested that, for golden eagles that fly within the calibrated maximum detection range for the species, operation of DTBird can be expected to reduce the overall likelihood of approaching the RSZ by 20–30%, with that estimate potentially further elevated to near 40% for birds at moderate to high Preexposure Risk of entering the RSZ.

The dwell time data could potentially be used as a surrogate for the pre-construction "eagle activity minutes" metric used to project fatality rates at wind-energy facilities using the Bayesian collision risk model developed by the U.S Fish and Wildlife Service. We could have independently compared projected post-construction fatality estimates tailored to the Goodnoe Hills based on dwell time at control turbines versus treatment turbines to create an estimate of fatality reduction. However, a comparison (# of fatalities/per year) of that scale could not be extrapolated to other facilities with different collision risk infrastructure and eagle activity rates and behaviors. Therefore, we determined a better approach was to present percentage estimates of DTBird's beneficial effects in reducing post-construction collision risk, which could potentially be tailored to match initial pre-construction facility projections tailored to specific sites using the USFWS Bayesian risk model. The results from the two study sites—one in a desert foothills landscape and one in temperate grassland ridgeline landscape—clearly indicated that DTBird's overall effectiveness may vary in different landscape/climatic settings with different resident and transient eagle populations, and variable false-positive deterrent-triggering rates that may influence the eagle responses. **A report for this milestone was submitted to DOE in January 2024 (Q27:M80) (Attachment 9).**

8.11 Task 11.0: Prepare systems cost analysis

8.11.1 Milestone 11.1 System cost analysis completed (Q28:M84)

When including the overall cost of Liquen´s Internal Services and R&D Department, the standard DTBirdV4D8 model sale cost (cameras model Falco and Larus software) is around $18K - $22K, and the yearly service sale cost around $2K - $3K. There are other project specific indirect costs for installation (around $4–6K/unit) and onsite maintenance (around $0.6–2K/unit/year). For the project, 16 DTBirdV4D8 units were manufactured in 2019 and delivered to Goodnoe Hills wind farm by the end of the year. Fourteen units operated under the evaluation and experiment from August 2021 to September 2023. **A report for this milestone is attached (Attachment 10)**

Table 2828. Actual Cost(s) to Install, Operate, and Maintain the DTBird system, Liquen ONLY (2016-2024).

Project Cost(s)	Amount (USD)	Unitary cost for the 14 units (USD)
ACTUAL DTBIRD PURCHASE COST FOR 14 UNITS	$208.619,64	$14.901,40

SHIPPING DTBIRDV4D8 UNITS TO GOODNOE HILLS SITE AND US CUSTOMS *	$17.114,49	$1.069,66
INSTALLATION COSTS (TRAVEL & SALARIES COSTS) – OCT 26TH TO NOV 3RD 2019	$10.659,23	$761,37
YEAR 1: TOTAL YEARLY SERVICE 13 DTBirdV4D8 (12 months) including technician travelling costs to repair multiple maintenance issues - August 2021 till July 2022	$42.997,43	$3.071,25
YEAR 2: TOTAL YEARLY SERVICE 14 DTBirdV4D8 (12 months) – August 2022 till September 2023	$35.199,41	$2.514,24
TOTAL 14 SYSTEMS + 24 MONTHS OF SERVICE	$327.278,51	$23.377,04

*16 units were delivered to the site

Section 9. Project Output/STI

9.1 Publications

No publications resulting from work performed under this Cooperative Agreement. Three draft manuscripts have been prepared to date, but they have not yet been submitted for publication.

9.2 Technologies/Techniques

No technologies or techniques were developed related to any aspect of the project with our knowledge under this Cooperative Agreement.

9.3 Status Reports

As part of the monthly check-in calls which REWI established with the DOE Contracting Team, unofficial status reports on this project were generated in advance of each monthly call. These status reports served as an agenda and guided the discussions during the calls and are preserved as attachments via email record between REWI and DOE (Attachment 11).

9.4 Media Reports

In November 2022, REWI mentioned “support from the U.S. Department of Energy, evaluation minimization technologies, including IdentiFlight and DTBird” in a REWI Special Update on the Eagle Rule. ***See*** ***https://rewi.org/2022/11/09/eagle-rule/.***

9.5 Invention Disclosures

No invention disclosures about any aspect of the project were made with our knowledge under this Cooperative Agreement.

9.6 Patent Applications

No patent applications related to any aspect of the project were submitted with our knowledge under this Cooperative Agreement.

9.7 Licensed Technologies

No subject inventions were licensed to third parties under this Cooperative Agreement.

9.8 Networks/Collaborations Fostered

No partnerships, networks or other means of collaboration were formed or concluded under this Cooperative Agreement.

9.9 Websites Featuring Project Work or Results

In September 2018, REWI (then AWWI) released "DTBird Technology Evaluation" and an accompanying technical report by H. T. Harvey & Associates detailing the results of an initial independent, site-specific pilot study of the DTBird detection/deterrence system at a wind facility in California. This initial study used UAVs and in-situ raptors to evaluate DTBird's ability to detect and deter large raptors, particularly golden eagles and reduce the risk of collisions with wind turbines. ***See*** ***https://rewi.org/resources/dtbird-technical-report*** **.**

9.10 Other Products

No additional project output was generated under this Cooperative Agreement.

9.11 Awards, Prizes, and Recognition

No awards or other forms of recognition were received by any party under this Cooperative Agreement.

Section 10. Project Summary Table

Task	Title / Task Description	Task Completion Date					Progress Notes
		Original Plan	Revised Plan (Mod 10)	Current Plan	Status	Percent Complete	

1.0	Project Launch and Development of Peer-Reviewed Study Design	Q1M3	Q5M15	Q5M15	Complete	100%	Task Completed
2.0	Expand Analysis of False Positives Using Data Collected During Pilot Study at California Wind Facility	Q4M9	Q5M15	Q5M15	Complete	100%	Task Completed
3.0	Evaluation of Pilot Study	Q4M10	Q6M18	Q6M18	Complete	100%	Task Completed
4.0	Update DTBird System and Revise Study Design for BP2 and BP3 As Appropriate	Q5M13	Q8M22	Q8M22	Complete	100%	Task Completed
5.0	Install DTBird at Goodnoe Hills, WA	Q5M15	Q18M53	Q18M53	Complete	100%	Task Completed
6.0	*In situ* CA, WA; UAV Trials WA	Q9M25	Q20M60	Q20M60	Complete	100%	Task Completed
7.0	Experiment Year 1	Q9M27	Q22M65	Q22M65	Complete	100%	Task Completed
8.0	Complete Controlled Experiment and Analyze Results	Q15M45	Q26M78	Q27:M79	Complete	100%	Task Completed
9.0	Evaluate Behavioral Responses of Raptors at Goodnoe Hills	Q15M45	Q27M79	Q25:M73	Complete	100%	Task Completed
10.0	Complete Combined Multi-Site Analyses	Q15M45	Q27M79	Q27:M80	Complete	100%	Task Completed
11.0	Systems Cost Analysis	Q15M45	Q28:M84	Q28:M84	Complete	100%	Task Completed
12.0	Final Report	Q15M45	Q28:M84	Q28:M84	Complete	100%	Task Completed